Probab. Theory Relat. Field 114, 309–399 (1999)

Probability Theory and Related Fields

Rescaled contact processes converge to super-Brownian motion in two or more dimensions

Richard Durrett[1,⋆], **Edwin A. Perkins**[2,⋆⋆]

[1] Department of Mathematics and ORIE, 278 Rhodes Hall, Cornell University, Ithaca, NY 14853, USA (e-mail: rtd1@cornell.edu)
[2] Department of Mathematics, 1984 Mathematics Rd., University of British Columbia, Vancouver, B.C. V6T 1Z2, Canada (e-mail: perkins@math.ubc.ca)

Received: 2 February 1998 / Revised version: 28 August 1998

Abstract. We show that in dimensions two or more a sequence of long range contact processes suitably rescaled in space and time converges to a super-Brownian motion with drift. As a consequence of this result we can improve the results of Bramson, Durrett, and Swindle (1989) by replacing their order of magnitude estimates of how close the critical value is to 1 with sharp asymptotics.

Mathematics Subject Classification (1991): Primary 60K35, 60G57; Secondary: 60F05, 60J80

1. Introduction

Our contact processes takes place on a fine lattice $\mathbf{Z}^d/M = \{z/M : z \in \mathbf{Z}^d\}$. The state of the process at time t is given by a function $\xi_t : \mathbf{Z}^d/M \to \{0, 1\}$, where $\xi_t(x) = 0$ indicates that x is vacant at time t and $\xi_t(x) = 1$ that the site is occupied by a particle. The dynamics of this right-continuous continuous time Markov chain can be described as follows:
(a) Particles die at rate 1 and give birth to one new particle at rate β.
(b) When a birth occurs at x the new particle is sent to a site y chosen at random from the $y \in \mathbf{Z}^d$ with $0 < \|y - x\|_\infty \le 1$, where $\|z\|_\infty = \sup_i |z_i|$.

⋆ Supported in part by NSF grant DMS-93-01070 and by an NSERC collaborative grant
⋆⋆ Supported in part by an NSERC research grant and an NSERC collaborative grant

(c) If y is vacant a new particle establishes itself there. If y is occupied, the birth is suppressed and no change occurs.

If $\beta < 1$ then particles die faster than they give birth and in addition lose births onto occupied sites, so the process dies out. To be precise, if we start with all sites occupied i.e., consider the process $\bar{\xi}_t$ starting from $\bar{\xi}_0(x) \equiv 1$ then the probability of an occupied site, $P(\bar{\xi}_t(x) = 1)$, which does not depend on x, tends to 0 as $t \to \infty$. Harris (1974) was the first to show that if β is large enough then (a) $P(\bar{\xi}_t(x) = 0)$ decreases to a positive limit as $t \to \infty$, and (b) the limit of the $\bar{\xi}_t$ defines a stationary distribution, $\bar{\xi}_\infty$. Simple monotonicity considerations tell us that conclusion (a) will hold for all β larger than

$$\beta_c = \inf\{\beta : P(\bar{\xi}_\infty(x) = 1) > 0\}$$

Self-duality of the contact process (see Theorem VI.1.7 of Liggett (1985)) shows that

$$\beta_c = \inf\{\beta : \lim_{t\to\infty} P(\xi_t \neq 0 | \xi_0 = \delta_0) > 0\} \ ,$$

i.e., β_c is also the critical birth rate for survival of the contact process starting with a single occupied site. Harris' original bound of β_c was very large but Holley and Liggett (1978) showed that in the nearest neighbor case that $\beta_c \le 4$ in all dimensions. There has been much work on numerical bounds for β_c in particular cases, most commonly the nearest neighbor one. See Stacey (1994) and Liggett (1985, 1995), but note that our parameter β is the total birth rate onto any site rather than than the rate of birth λ from a site to a particular neighbouring site. Chapter VI of Liggett (1985), and Durrett (1988), (1992b) are good places to learn about contact processes.

Bramson, Durrett, and Swindle (1989) considered the problem of the asymptotic behavior of the critical value $\beta_c(M)$ for the long range contact process as $M \to \infty$.

Theorem A. *As $M \to \infty$, $\beta_c(M) \to 1$. Furthermore*

$$\beta_c(M) - 1 \approx \begin{cases} C/M^{2/3} & d = 1 \\ C(\log M)/M^2 & d = 2 \\ C/M^d & d \ge 3 \end{cases}$$

where $\approx$ means if C is a small (large) positive number then the right hand side is a lower (upper) bound for large M.

To explain the answer, we begin by considering an inverse problem: given $\beta = 1 + \theta/N$, where $\theta > 0$, how large does M need to be to allow the contact process to survive? For the branching process with $M = \infty$

(so that new particles are never born onto occupied sites) the mean number of particles at time t is $\exp(t\theta/N)$, so the process will need time $O(N)$ to become significantly supercritical.

It is a well known fact that for the branching random walk when a typical particle at time N counts the number of its relatives within distance 1, the expected value of the result will be

$$\approx C \int_1^N t^{-d/2}\, dt = \begin{cases} CN^{1/2} & d = 1 \\ C \log N & d = 2 \\ C & d \geq 3 \end{cases} \tag{1.1}$$

Here $N - t$ is the time the nearby relative broke off from the ancestral tree of our typical particle and $Ct^{-d/2}$ is the probability that it stays near to the original particle. The excess birth rate above 1 is only θ/N in the contact process, so for this to compensate for the suppressed births we need M to be large enough so that the fraction of occupied sites will be of order $1/N$. That is, we choose M such that

$$M^d = \begin{cases} N^{3/2} & d = 1 \\ N \log N & d = 2 \\ N & d \geq 3 \end{cases} \tag{1.2}$$

Bramson, Durrett, and Swindle (1989) used branching process estimates and a block construction to show that with this choice of M the contact process dies out for small $\theta > 0$ and survives for large θ. To approach the problem of getting sharp asymptotics we will set $\beta = 1 + \theta/N$, compress time by a factor of N, and scale space by a factor of $N^{-1/2}$ to compensate for the time scaling. That is, we declare that:
(a) Particles die at rate N and give birth to one new particle at rate $N + \theta$.
(b) When a birth occurs at x the new particle is sent to a site y chosen at random from the $y \in \mathbf{Z}^d/(N^{1/2}M)$ with $0 < \|y - x\|_\infty \leq N^{-1/2}$.
(c) If y is vacant a new particle establishes itself there. If y is occupied, then the birth is suppressed and no change occurs.
Now θ is a fixed real number (although we are primarily interested in $\theta \geq 0$) and we only consider $N \in \mathbf{N}$ such that $N + \theta > 0$.

The case $d = 1$ has been previously studied by Mueller and Tribe (1995). To state their result, we rewrite the contact process as a set valued process by considering $\{x : \xi_t(x) = 1\}$ and consider the approximate density process

$$u_N(t, x) = \frac{1}{2N^{1/2}} \left| \xi_t \cap \left[x - N^{-1/2}, x + N^{-1/2}\right] \right|$$

To check this scaling note that for $d = 1$, $M = N^{3/2}$ so the lattice is $\mathbf{Z}/N^2$ and the above neighbourhood will contain $N^{3/2}$ sites but only about $O(N^{1/2})$ particles by (1.1) and our spatial scaling by $N^{-1/2}$. Let $\mathscr{C}_0$ denote the space of continuous functions from $\mathbf{R}$ to $[0, \infty)$ with compact support equipped with the topology of uniform convergence and let $\Omega_u = D([0, \infty), \mathscr{C}_0)$ be the Skorokhod space of cadlag $\mathscr{C}_0$-valued paths. A special case of their result shows

Theorem B. *If the initial conditions $u_N(0, x)$ approach $u(0, x)$ in $\mathscr{C}_0$ as $N \to \infty$, then $u_N(t, x)$ converges weakly in Ω_u to the solution of the stochastic partial differential equation*

$$du = \left(\frac{1}{6}u'' + \theta u - u^2\right)dt + \sqrt{2u}\, dW \tag{1.3}$$

Here u'' denotes differentiatioin with respect to x, dW is a space-time white noise (see Walsh (1986)), and we have considered a restricted class of initial conditions to avoid the issue of growth conditions at ∞. To explain the limit, the u'' results from displacement of particles with the $6 = 2 \cdot 3$ dictated by the fact that the uniform distribution on $[-1, 1]$ has variance $1/3$. The drift term θu comes from the "excess" birth rate, $-u^2$ reflects the lost births onto occupied sites, and the $\sqrt{2u}$ from the fact that we have births and deaths each at rate 1 per particle. To prepare for our discussion of our Theorem 1, note that in proving Theorem B, Mueller and Tribe showed tightness of the approximations in a space of continuous functions. Thus for large N, nearby sites have an almost identical number of occupied neighbors, i.e., the ratio of the number of neighbors at two nearby sites is close to 1.

Without the $-u^2$ the equation in (1.3) is the stochastic partial differential equation (SPDE) for the density of one-dimensional super-Brownian motion (see Reimers (1989) or Konno and Shiga (1988)). When $d \geq 2$, super-Brownian motion is singular with respect to Lebesgue measure so equations like (1.3) are meaningless. An alternate approach is to characterize super-Brownian motion as the solution of a measure-valued martingale problem. To this end we introduce the space $C_b^n(\mathbf{R}^d)$ of bounded continuous functions whose partial derivatives of order less than $n + 1$ are also bounded and continuous ($n \in \mathbf{N}$ or $n = \infty$). Let $M_F(\mathbf{R}^d)$ denote the space of finite measures on R^d with the topology of weak convergence, $\Omega_X = D([0, \infty), M_F(\mathbf{R}^d))$ be the Skorokhod space of cadlag $M_F(\mathbf{R}^d)$-valued paths, and $\Omega_{X,C}$ be the space of continuous $M_F(\mathbf{R}^d)$-valued paths with the topology of uniform convergence on compacts. Integration of a function ϕ with respect to a measure μ is denoted by $\mu(\phi)$. An adapted a.s.-continuous $M_F(\mathbf{R}^d)$-valued process $(X_t, t \geq 0)$ on a complete filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ ($\mathscr{F}_t$ is right-continuous) is an $(\mathscr{F}_t)$-super-Brownian motion starting at $X_0 \in M_F(\mathbf{R}^d)$ with branching rate $\gamma > 0$, diffusion coefficient $\sigma^2 > 0$ and drift $\theta \in \mathbf{R}$ if and only if it satisfies the following martingale problem:

$$(MP)^{\gamma,\sigma^2,\theta}_{X_0} \qquad \begin{aligned} &\text{For all } \phi \in C_b^\infty(\mathbf{R}^d), \\ &Z_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(\sigma^2\Delta\phi/2 + \theta\phi)ds \\ &\text{is an } (\mathscr{F}_t)\text{-martingale with } \langle Z(\phi)\rangle_t = \int_0^t X_s(\gamma\phi^2)ds \ . \end{aligned}$$

The law of X on $\Omega_{X,C}$ is then unique and $(MP)^{\gamma,\sigma^2,\theta}_{X_0}$ holds for a larger class of test functions including $C_b^2(\mathbf{R}^d)$. See Theorem 2.3 of Evans and Perkins (1994) for the latter. The uniqueness follows from Dawson's Girsanov theorem (see Theorem 5.1 of Dawson (1978)) and the uniqueness of the above martingale problem with $\theta = 0$. This latter result may be found in Dawson (1994) (Theorem 6.1.3) for a slightly larger class of test functions but as our class of functions contains a core for the generator of Brownian motion on the Banach space of continuous functions with limits at infinity (Ethier and Kurtz (1986), Proposition 5.1.1) uniqueness then follows in the above.

If instead of (1.2) we choose $M = \infty$, then the set of occupied sites at time t is a branching random walk $\zeta_t^N \subset \mathbf{R}^d$. Define a measure-valued process by

$$\mu_t^N(\phi) = \frac{1}{N}\sum_{x\in\zeta_t^N}\phi(x)$$

for all bounded measurable functions ϕ. Results in Chapter 4 of Dawson (1993) (see Theorem 4.6.2) then give

Theorem C. *If the initial measures μ_0^N approach μ_0 in $M_F(\mathbf{R}^d)$, then the sequence of measure-valued processes μ_t^N converges to super-Brownian motion μ_t starting at μ_0 with branching rate* 2, *diffusion coefficient* 1/3 *and drift* θ.

The purist may notice that our branching mechanism is slightly different than that in Dawson (1993) (only one particle jumps at each birth time) and Dawson works on the one-point compactification. The necessary modifications are straightforward, moreover Theorem C will also follow from the easy parts of our proof of Theorem 1 below.

Since the contact process can be dominated by the branching process that results by ignoring rule (c) and allowing births onto occupied sites, we must have a singular limit for the rescaled contact process in $d \geq 2$. We again assign each particle mass $1/N$ and look at the measure valued process defined by

$$X_t^N(\phi) = N^{-1} \sum_{x \in \xi_t} \phi(x)$$

Then $X_\cdot^N \in \Omega_X$. Here for reasons that will become clear in a moment we have suppressed the dependence on N in ξ_t. To state our limit result we need a definition. In $d \geq 3$ we let $u_1, u_2, \ldots$ be i.i.d. uniform on $[-1,1]^d$, define the random walk $U_n = u_1 + \cdots + u_n$, and let

$$b_d = \sum_{n=1}^{\infty} P\big(U_n \in [-1,1]^d\big)/2^d$$

In $d = 2$ we set $b_2 = 3/2\pi$.

Theorem 1. *Suppose that $d \geq 2$. If the initial measures X_0^N approach a limit X_0 in $M_F(\mathbf{R}^d)$ with no point masses, then the sequence of measure-valued processes $\{X_\cdot^N\}$ converges weakly on Ω_X to a super-Brownian motion $X_\cdot$ starting at X_0 with branching rate 2, diffusion coefficient $1/3$ and drift $\theta - b_d$.*

To explain Theorem 1, we begin with the easier case $d \geq 3$ and again look at the unscaled branching random walk in which births and deaths each occur at rates O(1). A closer look at the reasoning that led to (1.2) tells us that

$$C \int_L^N t^{-d/2}\, dt \sim C\, L^{1-d/2} \tag{1.4}$$

gives an upper bound on the expected number of neighbors y of a randomly chosen particle x at time N such that the last common ancestor of x and y was at a time before $N - L$. Here $x, y \in \mathbf{Z}^d/M_N$ are neighbors if $0 < \|x - y\|_\infty \leq 1$. This implies that if $|x_1 - x_2| \geq K_N$ where $K_N \to \infty$ then the number of neighbors of two particles x_1 and x_2 in the unit speed (and unscaled) branching random walk are "almost independent". To see this note that (1.4) shows that, up to a small error approaching 0 as $L \to \infty$, we only need consider contributions to the number of neighbors of x_1 and x_2 from cousins which branch off in the last time interval of length L. As $K_N \to \infty$ this shows that modulo a small error the number of neighbors of the two points depends on a disjoint set of random walk increments in the branching Brownian tree and so are "almost independent". Rescaling time and space we see that the number of neighbors of two particles x_1 and x_2 with $|x_1 - x_2| \geq K_N N^{-1/2}$ are almost independent. This and the law of large numbers implies that the amount of mass lost near a point x due to births onto an occupied site is just the mean number of neighbors of a randomly chosen point, b_d, times the mass there.

The reasoning described in the last paragraph just barely works in $d = 2$. Taking $L = N/\log N$ in (1.4) gives

$$\int_{N/\log N}^{N} t^{-1}\,dt \sim \log\log N = o(\log N)$$

So rescaling time and space, we see most of the neighbors of a particle are its relatives with most recent common ancestor less than $1/\log N$ back in time. Thus, if $K_N \to \infty$ the number of neighbors of two particles x_1 and x_2 with $|x_1 - x_2| \geq K_N(\log N)^{-1/2}$ are almost independent. Again the law of large numbers implies that the amount of mass lost near a point x due to births onto an occupied site is just a constant b_2 times the mass there, where

$$b_2 = \lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N/\log N} P\big(U_n \in [-1,1]^2\big)/2^2 = \frac{3}{2\pi}\ ,$$

the last by a local central limit theorem (see Section 8).

To turn the heuristics in the last two paragraphs into a proof, we will define a sequence of approximating processes ξ_t^k with the same initial condition ξ_0. Like the contact process, ξ_t, these depend on N but we will not exhibit the dependence in the notation. The first process in the sequence is the branching random walk ξ_t^0 which results if we ignore rule (c) in the definition of the contact process above and allow births onto occupied sites. Without the collision rule, ξ_t^0 may have more than one particle at a site so we regard ξ_t^0 as a "multi-set," i.e., a set in which repetitions of elements is allowed. For example, $\{a, a, b, b, b, c\}$ would represent two particles at a, three at b and one at c.

For $k \geq 1$ we let ξ_t^k be the branching random walk ξ_t^0 with the collision rule that births onto sites in ξ_t^{k-1} are suppressed. ξ_t^1 is an underestimate of the contact process ξ_t since it removes particles that collide with the larger set ξ_t^0. In the other direction ξ_t^2 is an overestimate since it removes only particles that collide with the smaller set ξ_t^1. The processes ξ_t^k are an alternating sequence of upper and lower bounds that, for fixed N and t, are equal to the contact process for $k \geq k_0(\omega, t)$, i.e., the number of iterations needed depends on the realization and the time of interest. We will not use this fact and so leave its verification to the interested reader. From the approximating processes ξ_t^k, $k \geq 0$ we can define measure-valued processes by

$$X_t^k(\phi) = N^{-1} \sum_{x\in\xi_t^k} \phi(x)$$

where sites are counted according to their multiplicities in ξ_t^k. In $d = 1$, we conjecture that as $N \to \infty$ these processes converge to limits that are all distinct. However, we can prove

Proposition 1. *Under the hypotheses of Theorem* 1, *if* $d \geq 2$, *then for all* $T > 0$,

$$E\left(\sup_{t \leq T} \left|X_t^2(1) - X_t^1(1)\right|\right) \to 0 \text{ as } N \to \infty$$

Note here that $X^2 \geq X^1$ so this result implies the total variation of $X_t^2 - X_t^1$ approaches 0 as $N \to \infty$ uniformly in $t \leq T$. Since the contact process X_t is trapped between X_t^2 and X_t^1 it follows that X_t is asymptotically the same as X_t^1. Given this, we can prove Theorem 1 by demonstrating

Proposition 2. *Under the hypotheses of Theorem* 1, *if* $d \geq 2$, *then* $X_\cdot^1$ *converges weakly in* Ω_X *to super-Brownian motion starting at* X_0 *with branching rate* 2, *diffusion coefficient* $1/3$ *and drift* $\theta - b_d$.

The process ξ_t^1 is much easier to analyze than the contact process since it is just the branching process minus particles that are born onto sites in the branching process. However, it still takes quite a bit of effort to prove Proposition 2. An outline of the proof can be found in Section 2. The details fill up Sections 3 to 10.

One reason for interest in Theorem 1 is that it allows us to sharpen the conclusion of a result of Bramson, Durrett, and Swindle (1989). Letting $V = (2M + 1)^d - 1$ be the number of neighbors a site has, and recalling that M and N are related by (1.2), we can now refine Theorem A as follows:

Theorem 2. *In* $d \geq 2$,

$$\beta_c(M) - 1 \sim \begin{cases} 2b_2(\log M)/M^2 \sim 4b_2(\log V)/V & \text{in } d = 2 \\ b_d/M^d \sim 2^d b_d/V & \text{in } d \geq 3 \end{cases},$$

where $\sim$ *means the ratio approaches one as* M (*or* V) *approaches* ∞.

The block construction, as described, for example, in Section 4 of Durrett (1995b), makes this a fairly straightforward consequence of Theorem 1. The details of the lower and upper bounds needed to prove Theorem 2 are given in Sections 11 and 12.

In $d = 1$, Mueller and Tribe (1994) have shown that the limiting SPDE in Theorem B has a critical value, θ_c, below which there is a.s. extinction and above which there is longterm survival with positive probability. In view of this it is natural to

Conjecture. *In $d = 1$ as $M \to \infty$,*

$$\beta_c(M) - 1 \sim \frac{\theta_c}{M^{2/3}}$$

To prove this seems difficult. Our proof of Theorem 2 makes crucial use of two facts: (i) the supercritical-subcritical phase transition in super-Brownian motion can be identified by by looking at the mean number of particles, and (ii) by computing second moments we can identify a suitable block event for the limiting process. Neither of these is available for the SPDE.

For $d = 2$, Theorem 2 gives the following asymptotic result for the critical value of the contact process for an L^∞ neighbourhood with V points:

$$\beta_c(V) - 1 \sim \frac{6}{\pi} \frac{\log V}{V} \tag{1.5}$$

To investigate the quality of this approximation for finite range M we have simulated the process with $M = 2$, i.e. $V = 25$.

For the simulation it is convenient to change the time scale so that $\beta = 1$ and the death rate, δ, is the parameter so that we can simulate the process for all values of δ simultaneously using the methods of Buttel, Cox and Durrett (1993). In the graph below we have plotted the fraction of occupied sites at time 5000 as a function of δ, starting from a configuration of all sites occupied in the 1000×1000 grid and with periodic boundary conditions. The estimate of β_c from the formula is about 1.25 which leads to an estimate of $4/5$ for the critical value of δ. This latter estimate is higher than the critical value of $2/3$ obtained from the simulation. Note, however, that the straight line from $(0, 1)$ to $(.8, 0)$ does a much better job of estimating the equilibrium density of particles than Theorem 2 of Bramson, Durrett and Swindle (1989) which, when rewritten in terms of δ, says

$$\lim_{N\to\infty} P\big(\bar{\xi}_\infty(x) = 1\big) = 1 - \delta \tag{1.6}$$

In principle one could use our methods to sharpen Theorem 3 of Bramson, Durrett and Swindle (1989) and give corrections to (1.6). We leave this task to an energetic reader.

2. Outline of the Proof

Choose $\theta \in \mathbf{R}$ and consider only $N \in \mathbf{N}$ with $N + \theta > 0$. We begin by constructing the processes which live on the rescaled lattices

$$\mathscr{L}_N = \begin{cases} N^{-1/2} \cdot N^{-1/d} \cdot \mathbf{Z}^d & d \geq 3 \\ N^{-1/2} \cdot N^{-1/2} (\log N)^{-1/2} \cdot \mathbf{Z}^2 & d = 2 \end{cases} \tag{2.1}$$

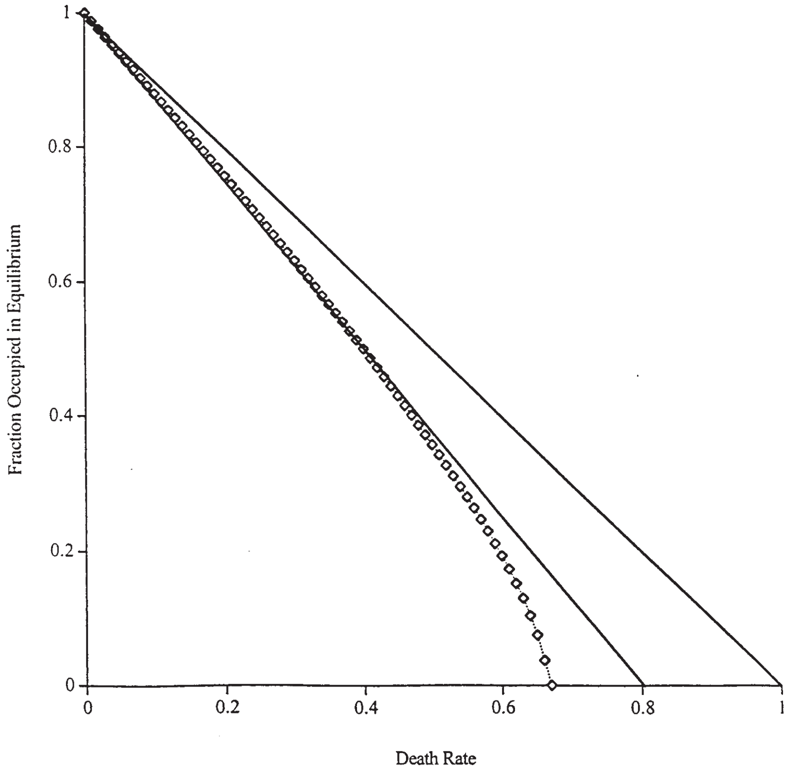

Figure 1

We have separated the scaling into two pieces since the first compensates for the fact that births occur at rate $N+\theta$, while the second increases the number of neighbors a site has. Since we want to relate the behavior of the long range contact process to that of a branching process, we will use a branching process type construction for the contact process rather than the usual graphical representation. Let

$$\mathscr{I} = \bigcup_{n=0}^{\infty} \left(\mathbf{N} \times \{0,1\}^n\right)$$

be the set of labels for the particles in our process. The first term in the union, $\mathbf{N} = \{1, 2, \ldots\}$, labels the initial particles $\{x_i^N : i \le K_N\} \subset \mathscr{L}_N$. Throughout, we will

ASSUME that the initial state consists of a finite number of particles that are located at distinct sites $\{x_i^N : i \le K_N\}$ and that $X_0^N = \frac{1}{N}\sum_{i=1}^{K_N} \delta_{x_i^N}$ approaches X_0 in $M_F(\mathbf{R}^d)$. (2.2)

To construct the time evolution we usually suppress dependence on N and work on a complete probability space $(\Omega, \mathscr{F}, P)$ containing the following independent collections of random variables:

$\{t_\beta : \beta \in \mathscr{I}\}$ are i.i.d. exponential with rate $2N + \theta$

$\{\delta_\beta : \beta \in \mathscr{I}\}$ are i.i.d. with $P(\delta_\beta = -1) = \dfrac{N}{2N+\theta}$

and $P(\delta_\beta = 1) = \dfrac{N+\theta}{2N+\theta}$

$\{e_\beta : \beta \in \mathscr{I}\}$ are i.i.d. with $P(e_\beta = 0) = P(e_\beta = 1) = 1/2$

$\{W_\beta : \beta \in \mathscr{I}\}$ are i.i.d. uniform on $\mathscr{N}_N = \left[-N^{-1/2}, N^{-1/2}\right]^d \cap \mathscr{Z}_N - \{0\}$

Deaths occur at rate N and births at rate $N + \theta$, so t_β is the time until a birth or death affects particle β. The event is a death if $\delta_\beta = -1$ and a birth if $\delta_\beta = 1$. The particle (β, e_β) is displaced from its parent β at B^β by an amount W^β, while the other particle $(\beta, 1 - e_\beta)$ remains at the location of β. (Here, (β, i) is a new member of $\mathscr{I}$ obtained from β by adding a coordinate $= i$ at the end.)

If $\beta = (\beta_0, \ldots, \beta_n) \in \mathscr{I}$ then we say β is in generation n and write $|\beta| = n$. If $m < |\beta|$ we write $\beta|m = (\beta_0, \ldots, \beta_m)$ for the ancestor of β in generation m. If $|\beta| > 0$ then we use $\pi\beta$ to denote its parent, i.e., its ancestor in generation $|\beta| - 1$. If $|\beta| = 0$ this is the -1 generation which does not exist, so we set $\pi\beta = \emptyset$, the empty string. From the definitions above it should be clear that

$$T_\beta = \sum_{m=0}^{|\beta|} t_{\beta|m}$$

is the end of the life of particle β. Since $\emptyset$ is a particle that doesn't exist, it is reasonable to declare that it died at time $-\infty$, i.e., $T_\emptyset = -\infty$. Let

$$\mathscr{F}_t = \sigma\left\{1(T_\alpha \le t)(T_\alpha, \delta_\alpha, e_\alpha, W_\alpha) : \alpha \in \mathscr{I}\right\}$$

which we assume has been completed by adding all the null sets.

To compute the positions of particles, we begin by noting that W_β is the displacement of the jumping particle at time T_β. The family line of β is moved when $\beta|m$ dies if $e_{\beta|m} = \beta_{m+1}$, so we define the position of the family line of β at time t by

$$\bar{B}_t^\beta = x_{\beta_0} + \sum_{m=0}^{|\beta|-1} W_{\beta|m}\, 1\left(e_{\beta|m} = \beta_{m+1}, T_{\beta|m} \le t\right) \tag{2.3}$$

The B here is meant to suggest Brownian motion, a stopped version of which is the limit of $\bar{B}_t^\beta$ if $N \to \infty$ and $|\beta|/N \to c$ (the stopping time).

The bar is added here so that it can be removed in the next paragraph in the displacement process that we will use repeatedly.

Up to this point we have ignored the deaths. To take them into account let

$$\zeta^0_\beta = T_\beta \wedge \inf \{T_{\beta|m} : \delta_{\beta|m} = -1, m \le |\beta|\} \tag{2.4}$$

Note that when the set of times is empty, we have $\inf \emptyset = \infty$ and $\zeta^0_\beta = T_\beta$. Since δ's produce deaths, the family line dies out at time ζ^0_β (or ceases to make sense at time T_β) and we let

$$B^\beta_t = \begin{cases} \bar{B}^\beta_t & t < \zeta^0_\beta \\ \Delta & t \ge \zeta^0_\beta \end{cases}$$

Here Δ is the usual cemetery state of Markov process theory used to indicate that the particle is no longer alive. In a number of instances we will be interested in the spatial location of particle β. Since β is alive on $[T_{\pi\beta}, T_\beta)$ and never moves, we can write

$$B^\beta_{T_{\pi\beta}} = B^\beta_{T_\beta -} \equiv B^\beta$$

where $\equiv$ indicates that the last quantity is a shorthand we will use for the first two.

We write $\beta \sim t$ if $T_{\pi\beta} \le t < T_\beta$. In words, $\beta \sim t$ if β labels a particle which *might* be alive at time t. Trivia buffs will want to note that if $|\beta| = 0$ then $\pi\beta = \emptyset$ and $T_\emptyset = -\infty$, so the particles in the initial configuration are actually alive at all negative times. We will adopt the convention that $\phi(\Delta) = 0$, for all functions ϕ so that dead particles don't contribute to our sums. Counting only the particles that are actually alive leads to our first measure-valued processes, the **branching random walk** X^0_t. This and all the subsequent measure-valued processes will be defined by integrating a bounded measurable test function ϕ with respect to the process:

$$X^0_t(\phi) = \frac{1}{N} \sum_{\beta \sim t} \phi(B^\beta_t) \tag{2.5}$$

Note that X^0_t depends on N even though we have not recorded this dependence in the notation. When we need to display the N, we will write $X^{0,N}_t$.

If μ is a measure, let $\mathrm{supp}(\mu)$ denote its closed support. If $t \to \mu_t$ is a measure-valued path which is cadlag, recall the definition of ζ^0_β from (2.4) and let

$$\zeta_\beta(\mu) = \zeta^0_\beta \wedge \inf\Big\{T_{\beta|m} : m < |\beta|,\ e_{\beta|m} = \beta_{m+1},\ B^\beta_{T_{\beta|m}} \in \mathrm{supp}\left(\mu_{T_{\beta|m}-}\right)\Big\}$$

In words the second term identifies the first time $T_{\beta|m}$ that a jump in the family line of β lands on a site that is already occupied, i.e., in μ_{t-}. With this notation introduced, we can define the **contact process** simply as the unique "strong solution" of

$$X_t(\phi) = \frac{1}{N}\sum_{\beta\sim t}\phi(B_t^\beta)1(\zeta_\beta(X) > t) \tag{2.6}$$

The existence and uniqueness of the solution are trivial for an initial finite set of particles since we can successively decide what to do at the event times T_α.

We can now define the sequence of processes $X^n,\ n \geq 1$ introduced in the previous section by:

$$X_t^n(\phi) = \frac{1}{N}\sum_{\beta\sim t}\phi(B_t^\beta)1\left(\zeta_\beta^n > t\right) \text{ where } \zeta_\beta^n = \zeta_\beta(X^{n-1}) \tag{2.7}$$

Note that in the case $n = 0$ this reduces to the definition of the branching random walk given in (2.5). (Anyone who is concerned that X^{n-1} is not defined when $n = 0$ can let this process be $\equiv 0$.)

Since $X_t^0 \geq X_t$ as measures (i.e., $X_t^0(\phi) \geq X_t(\phi)$ for all $\phi \geq 0$), comparing (2.6) and (2.7) shows $X_t^1 \leq X_t$, again as measures. Repeating this reasoning gives

$$X_t^1 \leq X_t \leq X_t^2 \leq X_t^0 \tag{2.8}$$

The first step in our derivation of Propositions 1 and 2 is to write down a stochastic equation for $X_t^n, n \geq 1$, which in the limit $N \to \infty$ will approach the martingale problem characterizing the limiting super-Brownian motion. We will only have to do this for $n = 1$ and $n = 2$ but for most of the proof it will be easier to write out the arguments for a general n. To derive our equation, we start with the observation that as t passes through time T_β we lose the particle β, but if we have a birth event then a new particle will exist at the same location, B^β, and a second particle will exist at $B^\beta + W_\beta$ if it does not land on an occupied site:

$$\begin{aligned} X_{T_\beta}^n(\phi) - X_{T_\beta-}^n(\phi) = \frac{1}{N}1(\zeta_\beta^n > T_{\pi\beta})\Big\{&-\phi(B^\beta) + 1(\delta_\beta = 1)\\ &[\phi(B^\beta) + 1\{B^\beta + W_\beta \notin \text{supp}(X_{T_\beta-}^{n-1})\}\\ &\times\phi(B^\beta + W_\beta)]\Big\}\end{aligned}$$

Taking advantage of the fact that $\delta_\beta \in \{1, -1\}$ we can rewrite the last expression as (just check the two cases)

$$\frac{1}{N}1(\zeta^n_\beta > T_{\pi\beta})\Bigg\{\phi(B^\beta)\delta_\beta + 1(\delta_\beta = 1)$$
$$\Big[\phi(B^\beta + W_\beta)1\{B^\beta_{T_{\pi\beta}} + W_\beta \notin \text{supp}(X^{n-1}_{T_\beta -})\} - \phi(B^\beta)\Big]\Bigg\} \tag{2.9}$$

Centering δ_β by subtracting its expected value, we can define $g_\beta = \delta_\beta - \theta/(2N+\theta)$ and split the first term (involving $\phi(B^\beta)\delta_\beta$) into two pieces:

$$\frac{1}{N}1\left(\zeta^n_\beta > T_{\pi\beta}\right)\phi\left(B^\beta\right)g_\beta + \frac{\theta}{N\,(2N+\theta)}1\left(\zeta^n_\beta > T_{\pi\beta}\right)\phi\left(B^\beta\right)$$

Define $h_\beta = 1(\delta_\beta = 1) - (N+\theta)/(2N+\theta)$ and do some arithmetic to write the second term in (2.9) as the sum of the following three terms:

$$\frac{1}{N}1(\zeta^n_\beta > T_{\pi\beta})h_\beta[\phi(B^\beta + W_\beta)1\{B^\beta_{T_{\pi\beta}} + W_\beta \notin \text{supp}(X^{n-1}_{T_\beta -})\} - \phi(B^\beta)]$$
$$+\frac{N+\theta}{N(2N+\theta)}1(\zeta^n_\beta > T_{\pi\beta})[\phi(B^\beta + W_\beta) - \phi(B^\beta)]$$
$$-\frac{N+\theta}{N(2N+\theta)}1(\zeta^n_\beta > T_{\pi\beta})[\phi(B^\beta + W_\beta)1\{B^\beta + W_\beta \in \text{supp}(X^{n-1}_{T_\beta -})\}]$$

To check this easily, begin by combining the second and third terms.

Summing in (2.9) over β with $T_\beta \le t$, telescoping the sum, recalling the above definition of h_β, and introducing

$$a^n_\beta(t) \equiv 1(T_\beta \le t, \zeta^n_\beta > T_{\pi\beta}) = 1(T_\beta \le t, \zeta^n_\beta \ge T_\beta), \quad n = 0, 1, 2, \ldots$$

as shorthand for "β was alive in X^n but died before time t," we have

$$\begin{aligned} X^n_t(\phi) = X^n_0(\phi) &+ \frac{1}{N}\sum_\beta a^n_\beta(t)\phi(B^\beta)g_\beta \\ &+ \theta\cdot\frac{1}{N(2N+\theta)}\sum_\beta a^n_\beta(t)\phi(B^\beta) + \frac{1}{N}\sum_\beta a^n_\beta(t)h_\beta \\ &\times[\phi(B^\beta + W_\beta)1\{B^\beta + W_\beta \notin \text{supp}(X^{n-1}_{T_\beta -})\} - \phi(B^\beta)] \\ &+ \frac{N+\theta}{N(2N+\theta)}\sum_\beta a^n_\beta(t)\left[\phi(B^\beta + W_\beta) - \phi(B^\beta)\right] \\ &- \frac{N+\theta}{N(2N+\theta)}\sum_\beta a^n_\beta(t)\phi(B^\beta + W_\beta)1\{B^\beta + W_\beta \in \text{supp}(X^{n-1}_{T_\beta -})\} \end{aligned} \tag{2.10}$$

By introducing notation for the various terms, we can rewrite the last equation briefly as

$$X_t^n(\phi) = X_0^n(\phi)+Z_t^n(\phi)+D_t^{n,1}(\phi)+E_t^{n,1}(\phi)+D_t^{n,2}(\phi)-K_t^n(\phi) \qquad (2.11)$$

Note this is valid for $n = 0$, in which case $K^0 \equiv 0$ by our convention that $X^{-1} \equiv 0$.

Here $K_t^n(\phi)$ is the "collision term" which counts the number of births onto occupied sites. The analysis of this term will be the hard part of the argument, so we begin with the other four terms. Here, and in what follows, $E_t^{n,i}(\phi)$ are "error" terms that will go to 0, the $D_t^{n,i}$ are "drift" terms that will have non-zero limits which are locally of bounded variation. Throughout this paper we will

$$\text{ASSUME that } \phi \in C_b^3 \quad \text{and} \quad \text{let } \|\phi\|_\infty = \max|\phi(x)| \ . \qquad (2.12)$$

For a number of the results this condition can be weakened to: ϕ is bounded and measurable (or ϕ: is Lipschitz continuous). However, we find it convenient to use one collection of test functions for all the results. Let $\epsilon_N = \theta/(2N+\theta)$. In Section 3 we establish the following results for $n \geq 1$:

Lemma 2.1. *$Z_t^n(\phi)$ is an $(\mathscr{F}_t)$-martingale with*

$$\langle Z^n(\phi)\rangle_t = \Big(2+\frac{\theta}{N}\Big)(1-\epsilon_N^2)\int_0^t X_r^n(\phi^2)dr$$

and $\langle Z^2(\phi) - Z^1(\phi)\rangle_t = \Big(2+\frac{\theta}{N}\Big)(1-\epsilon_N^2)\int_0^t (X_r^2(\phi^2) - X_r^1(\phi^2))dr.$

Lemma 2.2. *For all $t > 0$*

$$\lim_{N\to\infty} E\Big(\sup_{s\leq t}|D_s^{n,1}(\phi) - \theta\int_0^s X_r^n(\phi)\,dr|\Big) = 0$$

Lemma 2.3. *For all $t > 0$*

$$\lim_{N\to\infty} E\Big(\sup_{s\leq t}|D_s^{n,2}(\phi) - \int_0^s X_r^n(\Delta\phi/6)dr|\Big) = 0$$

Lemma 2.4. *For all* $t > 0$ $\lim_{N\to\infty} E\left(\sup_{s\leq t}|E_s^{n,1}(\phi)|\right) = 0$

The first three conclusions are straightforward to prove and are what one should guess by comparison with the corresponding parts of (2.10), i.e., the first, second, and fourth lines. To handle the error term $E_t^{n,1}(\phi)$, note that if we remove the mean-zero random variables h_β from the definition, it is equal to $D_t^{n,2}(\phi) - K_t^n(\phi)$. Intuitively, if $D_t^{n,2}(\phi) - K_t^n(\phi)$ stays bounded, then the h_β should cause cancellations that drive $E_t^{n,1}(\phi)$ to 0. To turn this idea into a proof we bound $K_t^n(\phi)$ above by the "collision term" of the branching random walk.

Lemma 2.5. *There is a constant* $0 < C < \infty$ *so that*

$$E\left(N^{-1}\sum_{\beta} a^0_\beta(t)1\{B^\beta + W_\beta \in \text{supp}(X^0_{T_\beta -})\}\right) \le C(X^0_0(1) + X^0_0(1)^2)$$

Notice about constants. This result has the first of a large number of C's that will appear. All of these constants may depend on the drift θ, the time t (or s), and (though it is vacuous here) on the test function ϕ. However, for fixed ϕ our constant $C(\theta, t)$ will be bounded on compact subsets of $\mathbf{R}\times[0,\infty)$. This condition is obviously satisfied whenever $C_\phi(\theta, t)$ is a continuous function. C will never depend on N, or on the initial condition X^N_0.

Lemma 2.5 indicates that we have enough neighbors so that the amount of mass lost due to interference is $O(1)$. This result is proved in Section 4, by proving two results, Lemmas 4.1 and 4.4 that investigate the amount of interference between (i) unrelated individuals and (ii) individuals with a common ancestor in generation 0. In Lemmas 5.1 and 5.3, we sharpen the two bounds in Section 4 to show that if the initial conditions $X^{0,N}_0 \Rightarrow X_0$ a measure with no point masses, then "collisions between distant relatives can be ignored." Here distant means that their most recent common ancestor was more than τ_N units of time in the past where

(i) in $d > 2$, $N\tau_N \to \infty$ and $\tau_N \to 0$.

(ii) in $d = 2$, $\tau_N = 1/\log N$.

These observations are the key to our result for the collision term.

Lemma 2.6. *Then for* $n = 1, 2$ *and any* $t < \infty$

$$\lim_{N\to\infty} E\left(\sup_{s\le t}\left|K^n_s(\phi) - \int_0^s b_d X^1_r(\phi)dr\right|\right) = 0$$

To argue intuitively, note that if C_N is large then two individuals that are related within τ_N units of time lie within $C_N\sqrt{\tau_N}$ distance in space with high probability. From this we see that the collision term has a correlation length that vanishes in the limit, and the number of collisions in a region becomes a constant times the mass of the process there.

An outline of the proof of Lemma 2.6 can be found in Section 6. There the result is broken down into seven lemmas that are proved in Sections 6–10. Rather than describe those technicalities now, we will instead explain why Lemmas 2.1–2.6 are enough to prove Propositions 1 and 2, and hence Theorem 1. The next step in that direction is the

Proof of Proposition 1. Subtracting (2.11) with $\phi = 1$ and $n = 2$ from the same formula with $n = 1$ and using Lemmas 2.1–2.6 shows that

$$X_t^1(1) - X_t^2(1) = Z_t^1(1) - Z_t^2(1) + (\theta - b_d) \int_0^t X_r^1(1) - X_r^2(1) dr + \hat{E}_t \quad (2.13)$$

where $u_N(t) = E(\sup_{s \le t} |\hat{E}_s|)$ satisfies $u_N(t) \to 0$. Since $Z_t^1(1) - Z_t^2(1)$ is a martingale, taking expected values and letting

$$f_N(t) = E[X_t^1(1) - X_t^2(1)] \ge 0$$

shows that

$$f_N(t) \le \theta^+ \int_0^t f_N(r) dr + u_N(t)$$

Let $F_N(t) = e^{\theta^+ t} u_N(t)$. A form of Gronwall's lemma implies that

$$\sup_{s \le t} f_N(s) \le \sup_{s \le t} F_N(s) = F_N(t)$$

and hence

$$\sup_{s \le t} E(X_s^1(1) - X_s^2(1)) \le F_N(t) \to 0 \quad (2.14)$$

To put the supremum inside the expected value we have to look at the martingale difference $M_t = Z_t^1(1) - Z_t^2(1)$. The L^2 maximal inequality (applied to M_t) combined with the fact that $M_t^2 - \langle M \rangle_t$ is a martingale null at 0 implies that

$$E\left(\sup_{s \le t} |Z_s^1(1) - Z_s^2(1)|^2 \right) \le CE|Z_t^1(1) - Z_t^2(1)|^2 = CE\langle Z^1(1) - Z^2(1) \rangle_t \quad (2.15)$$

Using Lemma 2.1 now, we have

$$\langle Z^1(1) - Z^2(1) \rangle_t = \left(2 + \frac{\theta}{N}\right)(1 - \epsilon_N^2) \int_0^t (X_r^1(1) - X_r^2(1)) dr \quad (2.16)$$

Combining (2.14)–(2.16) we have that as $N \to \infty$

$$v_N(t) \equiv E\left(\sup_{s \le t} |Z_s^1(1) - Z_s^2(1)| \right) \to 0 \quad (2.17)$$

Returning to (2.13) now, we can let $g_N(t) \equiv E\big(\sup_{s \le t} |X_s^1(1) - X_s^2(1)|\big)$, note that the difference inside the absolute values is always positive, and conclude

$$g_N(t) \le v_N(t) + \theta^+ \int_0^t g_N(r)dr + u_N(t)$$

Again if we let $G_N(t) = e^{\theta^+ t}(u_N(t) + v_N(t))$ then $g_N(t) \le G_N(t) \to 0$. □

Turning now to Proposition 2, we have from the above lemmas:

For each $\phi \in C_b^3(\mathbf{R}^d)$
$X_t^1(\phi) = X_0^1(\phi) + Z_t^1(\phi) + \int_0^t X_s^1((\theta - b_d)\phi + \Delta\phi/6)ds + \hat{E}_t^1(\phi)$
$Z_t^1(\phi)$ is an $\mathscr{F}_t$-martingale as in Lemma 2.1 and
$\lim_{N\to\infty} E(\sup_{s\le t} |\hat{E}_s^1(\phi)|) = 0$. (2.18)

Let P_N denote the law of X^1 on Ω_X and $X_t(\omega) = \omega(t)$ denote the coordinate variables on Ω_X. To prove tightness of $\{P_N\}$ we use the following specialized version of Jakubowski's general criterion on $D([0,\infty), E)$ for E Polish (see Theorem 3.6.4 of Dawson (1993)). Recall that $\Phi \subset C_b(\mathbf{R}^d)$ is a separating class iff the integrals $\{\mu(\phi) : \phi \in \Phi\}$ uniquely determine μ in $M_F(\mathbf{R}^d)$.

Lemma 2.7. *Let Φ be a separating class which is closed under addition. A sequence of probabilities $\{\tilde{P}_N\}$ on Ω_X is tight iff the following conditions hold*:

(i) *For each $T, \epsilon > 0$, there is a compact set $K_{T,\epsilon} \subset \mathbf{R}^d$ such that*

$$\sup_N \tilde{P}_N\left(\sup_{t\le T} X_t(K^c_{T,\epsilon}) > \epsilon\right) < \epsilon \ .$$

(ii) $\lim_{M\to\infty} \sup_N \tilde{P}_N(\sup_{t\le T} X_t(1) > M) = 0$.

(iii) *If $\tilde{P}_N^\phi(A) = \tilde{P}_N(X_\cdot(\phi) \in A)$, then for each $\phi \in \Phi$, $\{\tilde{P}_N^\phi : N \in \mathbf{N}\}$ is tight in $D = D([0,\infty), \mathbf{R})$.*

The derivation of this result from the more general results cited above is straightforward (see Theorem 3.7.1 of Dawson (1993) for the slightly simpler setting of the one-point compactification of $\mathbf{R}^d$).

Recall that $\Omega_{X,C}$ is the space of continuous $M_F(\mathbf{R}^d)$-valued paths with the compact-open topology. Specializing the above result further we have

Lemma 2.8. *Assume $\{\tilde{P}_N\}$ satisfy hypothesis* (i) *of Lemma* 2.7 *and for each ϕ in $C_b^\infty(\mathbf{R}^d)$, $\{\tilde{P}_N^\phi : N \in \mathbf{N}\}$ is tight in D and all limit points are supported by $C = C([0,\infty), \mathbf{R}^d)$. Then $\{\tilde{P}_N\}$ is tight in Ω_X and all limit points are supported on $\Omega_{X,C}$.*

Proof. Taking $\phi = 1$, we see that our assumption on $\{\tilde{P}_N^1\}$ readily implies (ii) in Lemma 2.7 (see Theorem 3.10.2 of Ethier and Kurtz (1986)). Lemma

2.7 shows that $\{\tilde{P}_N\}$ is tight on Ω_X. Let $\tilde{P}$ be a limit point. If $\phi \in C_b^\infty(\mathbf{R}^d)$ then $\tilde{P}^\phi(\cdot) = \tilde{P}(X(\phi) \in \cdot)$ is a limit point of $\{\tilde{P}_N^\phi\}$ and so is supported by C. Let Φ_0 be a countable subset of the functions in C_b^∞ with compact support which is dense in the space of continuous functions with compact support in $\mathbf{R}^d$. Then $X_\cdot(\phi)$ is continuous for all $\phi \in \Phi_0$ $\tilde{P} - a.s.$ As Φ_0 is a separating class, this implies $X_t = X_{t-}$ for all $t \geq 0$ $\tilde{P} - a.s.$ □

Lemma 2.9. *Let P_x^N denote the law of a continuous time random walk B_t starting at x which at rate $N + \theta$ takes a step uniformly distributed over $\mathcal{N}_N$. If $\phi : \mathbf{R}^d \to \mathbf{R}$ is bounded and measurable then*

$$E(X_t^0(\phi)) = e^{\theta t} \int E_x^N(\phi(B_t)) X_0^0(dx)$$

and there is a constant $C = C(\theta, t)$ such that

$$E(X_t^0(1)^4) \leq C[X_0^0(1) + X_0^0(1)^4] \ .$$

Proof. This follows from the known moment measures of a branching random walk starting from a single particle. See for example Lemma 2.2 of Bramson, Durrett and Swindle (1989) (and set $\lambda = 1+(\theta/N)$ in that result). A few simple moment inequalities for sums of i.i.d. random variables are needed to derive the second result from the lemma in Bramson et al. which assumes a single initial particle. □

Lemma 2.10. *$\{P_N\}$ is tight on Ω_X and all limit points are supported by $\Omega_{X,C}$.*

Proof. We apply Lemma 2.8. Let $\epsilon \in (0, 1)$, $T > 0$, and for $R > 1$ choose a C^∞ function $h_R : \mathbf{R}^d \to [0, 1]$ such that

$$B(0, R) \subset \{x : h_R(x) = 0\} \subset \{x : h_R(x) < 1\} \subset B(0, R+1)$$

and all the derivatives of h_R of order two or less are uniformly bounded in $(x, r) \in \mathbf{R}^d \times (1, \infty)$. By (2.2), Lemma 2.9 and the weak convergence of the random walks in that result to Brownian motion, we may choose R sufficiently large so that

$$\sup_N \sup_{t \leq T} E(X_t^0(B(0, R)^c)) < \epsilon^3 \tag{2.19}$$

The analogue of (2.18) for X_t^0 (omit the killing term) gives:

$$X_t^0(h_R) = X_0^0(h_R) + Z_t^0(h_R) + \int_0^t X_s^0(\theta h_R + \Delta h_R/6) ds + \hat{E}_t^0(h_R) \ ,$$

where $Z^0_t(h_R)$ is an $(\mathscr{F}_t)-$ martingale such that

$$\langle Z^0(h_R)\rangle_t = \left(2+\frac{\theta}{N}\right)(1-\epsilon_N^2)\int_0^t X_s^0(h_R^2)ds$$

and

$$\lim_{N\to\infty} E(\sup_{t\le T}|\hat{E}^0_t(h_R)|) = 0 \text{ for all } T>0$$

Therefore, applying Chebychev on each term and then (2.19), we see that

$$\begin{aligned}
P\left(\sup_{t\le T} X_t^0(h_R) > 4\epsilon\right) &\le 1(X_0^0(h_R)>\epsilon) + \frac{c}{\epsilon^2}\int_0^T E(X_s^0(h_R^2))ds \\
&\quad + \frac{c}{\epsilon}\int_0^T E(X_s^0(B(0,R)^c))ds \\
&\quad + \frac{1}{\epsilon}E\left(\sup_{t\le T}|\hat{E}^0_t(h_R)|\right) \\
&\le cT\epsilon + \frac{1}{\epsilon}E\left(\sup_{t\le T}|\hat{E}^0_t(h_R)|\right) \\
&\le (cT+1)\epsilon \quad ,
\end{aligned}$$

where the last inequality is valid for N sufficiently large. As $X^1 \le X^0$, (i) of Lemma 2.7 is now obvious.

Fix $\phi \in C_b^\infty(\mathbf{R}^d)$. We will use (2.18) to verify the other hypothesis of Lemma 2.8. If $0\le t<u\le T$, then

$$\begin{aligned}
E\left(\left(\int_t^u X_s^1((\theta-b_d)\phi+\Delta\phi/6)ds\right)^2\right) &\le c(\phi)E\left(\left(\int_t^u X_s^1(1)ds\right)^2\right) \\
&\le c(\phi)\sup_{s\le T}E(X_s^0(1)^2)(u-t)^2 \le c(\phi)(u-t)^2 \quad ,
\end{aligned}$$

where we have used Lemma 2.9 in the last line. This shows that $C_N(t) = \int_0^t X_s^1((\theta-b_d)\phi+\Delta\phi/6)ds$ defines a tight sequence of processes on C (e.g. by Theorem 8.3 of Billingsley (1968)).

Turning now to the martingale terms in (2.18), arguing as above we see from Lemma 2.1 that $\{\langle Z^1(\phi)\rangle_t : N\in\mathbf{N}\}$ is a tight sequence of processes in C. Note that by definition, $\sup_{t\le T}|\Delta Z_t^1(\phi)| \le 2\|\phi\|_\infty N^{-1}$. Theorem VI.4.13 and Proposition VI.3.26 of Jacod and Shiryaev (1987) now show that $\{Z^1_\cdot(\phi) : N\in\mathbf{N}\}$ is a tight sequence in D and all limit points are supported by C. These results with (2.18) and Corollary VI.3.33 of Jacod and Shiryaev (1987) show that $\{P_N^\phi : N\in\mathbf{N}\}$ is tight in D and that all limit points are supported on C. Lemma 2.8 now completes the proof. □

It is now straightforward to prove Proposition 2. Let P be a limit point of $\{P_N\}$. Then P is a law on $\Omega_{X,C}$ and we must show it satisfies the martingale problem $(MP)_{X_0}^{2,1/3,\theta-b_d}$ from Section 1 which characterizes super-Brownian motion with the appropriate parameters. By Skorokhod's theorem we may assume that (now making dependence on N explicit) $X^{1,N_k} \to X^1$ in Ω_X a.s. Let $\phi \in C_b^\infty(\mathbf{R}^d)$ and set

$$Z_t(\phi) = X_t^1(\phi) - X_0^1(\phi) - \int_0^t X_s^1((\theta - b_d)\phi + \Delta\phi/6)ds$$

We must show that $Z_t(\phi)$ and $Z_t(\phi)^2 - \int_0^t X_s^1(2\phi^2)ds$ are $(\mathscr{F}_t^{X^1})$-martingales under P, where $(\mathscr{F}_t^{X^1})$ is the canonical right-continuous filtration generated by X^1. We only show the latter as it is slightly more involved. Fix $0 \le t_1 < \cdots < t_n \le s < t$, and let $h_i : M_F(\mathbf{R}^d) \to \mathbf{R}$ be bounded and continuous for $i \le n$. Write $Z_t^{1,N_k}(\phi)$ for the martingale term in (2.18) with $N = N_k$. By taking another subsequence we see from (2.18) that $\sup_{t \le T} |Z_t^{1,N_k}(\phi) - Z_t(\phi)| \to 0$ for all $T > 0$ a.s. Use Lemma 2.9 for the necessary uniform integrability to conclude

$$\begin{aligned}
&E\left(\left(Z_t(\phi)^2 - Z_s(\phi)^2 - \int_s^t X_r^1(2\phi^2)\,dr\right)\prod_1^n h_i(X_{t_i}^1)\right)\\
&= \lim_{k\to\infty} E\Bigg(\Bigg(Z_t^{1,N_k}(\phi)^2 - Z_s^{1,N_k}(\phi)^2\\
&\quad - \int_s^t X_r^{1,N_k}\left(\left(2 + \frac{\theta}{N_k}\right)\left(1 - \epsilon_{N_k}^2\right)\phi^2\right)dr\Bigg)\prod_1^n h_i\left(X_{t_i}^{i,N_k}\right)\Bigg)\\
&= 0 \text{ (by (2.18))}
\end{aligned}$$

This completes the proof of Proposition 2 and hence proves Theorem 1, modulo Lemmas 2.1–2.6.

3. The four easy convergences

In this section, we have two aims. First we will introduce some useful martingales. Then we will prove Lemmas 2.1–2.4.

Lemma 3.1. (a) *$\mathscr{F}_t$ is a right-continuous filtration.*
(b) *For each $\beta \in \mathscr{I}$, B_t^β is $\mathscr{F}_t$-optional.*
(c) *For all $\beta \in \mathscr{I}$ and $n \ge 0$, T_β, and ζ_β^n are $\mathscr{F}_t$-stopping times.*
(d) *For all $n \ge 0$, X^n and X are $\mathscr{F}_t$-optional.*

These claims are intuitively obvious and formal proofs are not hard to construct so we proceed to:

Lemma 3.2. *Let $\psi : [0, \infty) \times \Omega \to \mathbf{R}$ be bounded and $\mathscr{F}_t$-predictable and let $\beta \in \mathscr{I}$ with $|\beta| > 0$. Then the following process is an $\mathscr{F}_t$-martingale*:

$$\psi(T_\beta, \omega) 1(T_\beta \le t) - (2N + \theta) \int_0^t 1(T_{\pi\beta} < r \le T_\beta) \psi(r, \omega) dr$$

Proof. Let N_t be the number of arrivals by time t in a rate λ Poisson process, and let $T_n = \inf\{t : N_t = n\}$ be the time of the nth arrival. As is well known $M_t = N_t - \lambda t$ is a martingale with respect to $\mathscr{G}_t = \sigma(T_n 1(T_n \le t) : n \in \mathbf{N})$, and so is the stochastic integral

$$\begin{aligned}
&\int_0^t \psi(r, \omega) 1(T_{n-1} < r \le T_n) dM_r \\
&\quad = \psi(T_n, \omega) 1(T_n \le t) - \lambda \int_0^t 1(T_{n-1} < r \le T_n) \psi(r, \omega) dr
\end{aligned}$$

If we take $\lambda = 2N + \theta$, $n = |\beta| + 1$ we obtain the desired result for the filtration $\mathscr{F}_t^0 = \sigma(T_{\beta|i} 1(T_{\beta|i} \le t) : i \le |\beta|)$ and hence also for the larger filtration obtained by adjoining the independent information in

$$\sigma(t_\gamma : \gamma \text{ not an ancestor of } \beta) \vee \sigma(\delta_\gamma, e_\gamma, W_\gamma : \gamma \in I)$$

As this is larger than $\mathscr{F}_t$, the result follows. □

Our second class of martingales is

Lemma 3.3. *Let ϕ be bounded and measurable, let $\beta \in \mathscr{I}$ and recall $g_\beta = \delta_\beta - \epsilon_N$ where $\epsilon_N = \theta/(2N + \theta)$. Then the following process is an $\mathscr{F}_t$-martingale*:

$$J(t) = 1\{T_\beta \le t, T_{\pi\beta} < \zeta_\beta^n\} \phi(B^\beta) g_\beta$$

Proof. Recall that $\mathscr{F}_{T_\beta -}$ is generated by $\{B \cap \{T_\beta > t\} : B \in \mathscr{F}_t, t \ge 0\}$, a class of sets closed under finite intersections. It is then straightforward to check that

$$E(J(T_\beta) | \mathscr{F}_{T_\beta -}) = 1\{T_\beta \le t, T_{\pi\beta} < \zeta_\beta^n\} \phi(B^\beta) E(g_\beta | \mathscr{F}_{T_\beta -}) = 0$$

Since $t \to J(t)$ is constant except for a jump at T_β, equal to $J(T_\beta)$ it follows that $J(t)$ is an $\mathscr{F}_t$-martingale. □

Before extending the class of martingales in Lemma 3.3, we need a technical result which will enable us to compute or bound various random variables. Let $r(\theta, t) = \int_0^t e^{\theta s}\, ds$.

Lemma 3.4. (a) $E\Big[\Big(\sum_\beta 1(T_\beta \le t)\Big)^p\Big] < \infty$ *for any* $0 < p < \infty$

(b) $E(N^{-1}\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta_\beta^0)) = (2N+\theta)r(\theta,t)X_0^0(1)$

Proof. To prove (a) consider Y_t, the number of particles that would contribute to our branching random walk X_t^0 if we start with $NX_0^0(1)$ individuals and have no deaths. That is, at each event T_β we ignore the δ_β's so a new particle is born and the old one does not die. Clearly, $\sum_\beta 1(T_\beta \le t) \le Y_t$. Since Y_t is a branching process in which each particle gives birth at rate $2N+\theta$, it is a speeded up version of the Yule process. It has long been known, see e.g., Kendall (1949), that Y_t has a geometric distribution, so $EY_t^p < \infty$ for all $p > 0$.

To prove (b) we begin by observing that

$$X_{T_\beta}^0(1) - X_{T_\beta-}^0(1) = N^{-1}1(T_\beta \le \zeta_\beta^0)\delta_\beta$$

Recalling $g_\beta = \delta_\beta - \theta/(2N+\theta)$ and $a_\beta^0(t) = 1(T_\beta \le t, T_\beta \le \zeta_\beta^0)$, multiplying the above equation by $1(T_\beta \le t)$, and summing over β, which is legitimate because of (a), we have

$$X_t^0(1) - X_0^0(1) = \frac{1}{N}\sum_\beta a_\beta^0(t)g_\beta + \frac{\theta}{N(2N+\theta)}\sum_\beta a_\beta^0(t)$$

The first term on the right-hand side is a martingale by Lemma 3.3 (part (a) proved above justifies integrability). Taking the expected value and using Lemma 2.9 gives for $\theta \ne 0$,

$$E\left(N^{-1}\sum_\beta a_\beta^0(t)\right) = \frac{2N+\theta}{\theta}E\left(X_t^0(1) - X_0^0(1)\right) = (2N+\theta)r(\theta,t)X_0^0(1)$$

The result is now also immediate for $\theta = 0$ by monotonicity in θ of the left-hand side of (b) and continuity in θ of the right-hand side. □

Lemma 3.5. *Assume that* G_β *is measurable with respect to* $\mathscr{F}_{T_\beta}$, $E(G_\beta|\mathscr{F}_{T_\beta-}) = 0$, *and*

$$|G_\beta| \le K1(T_\beta \le \zeta_\beta^0)$$

Then $M_t = N^{-1}\sum_\beta 1(T_\beta \le t)G_\beta$ *is an* $\mathscr{F}_t$*-martingale and there is a* $0 < C < \infty$ *so that*

$$E\left(\sup_{s\le t} M_s^2\right) \le CK^2X_0^0(1)$$

Proof. Let $J_\beta(t) = 1(T_\beta \le t)G_\beta$. With our assumptions we can follow the proof of Lemma 3.3 to conclude that $J_\beta(t)$ is a martingale. Summing up these martingales with (a) of Lemma 3.4 to check integrability shows that M_t is a martingale.

To prove the bound we recall that since our martingale has paths of bounded variation, $[M]_t = \sum_{s\le t}(M_s - M_{s-})^2$. Using the L^2 maximal inequality for martingales with the fact that $M_t^2 - [M]_t$ is a martingale that is null at 0 gives

$$E\left(\sup_{s\le t} M_s^2\right) \le C E[M]_t \tag{3.1}$$

Combining this with the formula for $[M]_t$ we have

$$\begin{aligned} E\left(\sup_{s\le t} M_s^2\right) &\le \frac{C}{N^2} E \sum_\beta 1(T_\beta \le t) G_\beta^2 \\ &\le \frac{CK^2}{N^2} \cdot E \sum_\beta 1(T_\beta \le t, T_\beta \le \zeta_\beta^0) \\ &\le CK^2 \cdot \frac{2N+\theta}{N} r(\theta, t) X_0^0(1) \end{aligned}$$

by $|G_\beta| \le K1(T_\beta \le \zeta_\beta^0)$ and (b) in Lemma 3.4. The result is now immediate by our convention on constants. □

Recall from Section 2, that $Z_t^n(\phi) = \frac{1}{N}\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta_\beta^n)\phi(B^\beta)g_\beta$ and we have assumed $\phi \in C_b^3(\mathbf{R}^d)$, although the next result only requires ϕ to be bounded and measurable. Lemmas 2.1–2.4 from Section 2 will now be proved. We restate them for the reader's convenience.

Lemma 2.1. *$Z_t^n(\phi)$ is an $\mathscr{F}_t$-martingale with*

$$\langle Z^n(\phi)\rangle_t = (1-\epsilon_N^2)\left(2+\frac{\theta}{N}\right)\int_0^t X_r^n(\phi^2)dr \quad ,$$

and

$$\langle Z^2(\phi) - Z^1(\phi)\rangle_t = (1-\epsilon_N^2)\left(2+\frac{\theta}{N}\right)\int_0^t X_r^2(\phi^2) - X_r^1(\phi^2)dr \quad .$$

Proof. The fact that $Z_t^n(\phi)$ is an $\mathscr{F}_t$-martingale follows from Lemmas 3.3 and 3.4(a) (the latter for the necessary integrability). Clearly,

$$[Z^n(\phi)]_t = N^{-2}\sum_\beta 1\left(T_\beta \le t, T_\beta \le \zeta_\beta^n\right)\phi^2(B^\beta)g_\beta^2$$

To convert this into $\langle Z^n(\phi)\rangle_t$, we will replace g_β^2 by its mean and then the sum by its compensator. Recalling various definitions we see that

$$Eg_\beta^2 = \mathrm{var}(\delta_\beta) = 1 - (E\delta_\beta)^2 = 1 - \epsilon_N^2$$

As in Lemma 3.3 one may readily check that $E(g_\beta^2|\mathscr{F}_{T_\beta-}) = 1 - \epsilon_N^2$ and so an application of Lemma 3.5 implies the following is a martingale:

$$M_t = N^{-2}\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta_\beta^n)\phi^2(B^\beta)\{g_\beta^2 - (1-\epsilon_N^2)\}$$

Applying Lemma 3.2 (and Lemma 3.4 (a)) with $\psi(r,\omega) = 1(r \le \zeta_\beta^n)\phi(B_{r-}^\beta)^2$ we see that

$$\begin{aligned} N_t = &\frac{1}{N^2}\sum_\beta 1\{T_\beta \le t, T_\beta \le \zeta_\beta^n\}\phi^2(B^\beta) \\ &- \frac{2N+\theta}{N}\int_0^t \frac{1}{N}\sum_\beta 1\{T_{\pi\beta} < r < T_\beta, r < \zeta_\beta^n\}\phi^2(B_r^\beta)dr \end{aligned}$$

is a martingale. Recalling the definition of $X_r^n(\phi^2)$ and using the fact that $M_t + (1-\epsilon_N^2)N_t$ is a martingale we have shown

$$\langle Z^n(\phi)\rangle_t = (1-\epsilon_N^2)\left(2 + \frac{\theta}{N}\right)\int_0^t X_r^n(\phi^2)dr$$

For the second assertion note that

$$Z_t^2(\phi) - Z_t^1(\phi) = \frac{1}{N}\sum_\beta 1(T_\beta \le t)1(\zeta_\beta^1 < T_\beta \le \zeta_\beta^2)g_\beta\phi(B^\beta)$$

and make minor changes in the above argument. □

We next consider: $D_t^{n,1}(\phi) = \theta\cdot\frac{1}{N(2N+\theta)}\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta_\beta^n)\phi(B^\beta)$. The following result again only requires ϕ to bounded and measurable. We will need this generalization for the proof of Lemma 2.3.

Lemma 2.2. *For all* $t > 0 \lim_{N\to\infty} E\left(\sup_{s\le t}|D_s^{n,1}(\phi) - \theta\int_0^s X_r^n(\phi)dr|\right) = 0$

Proof. Lemma 3.2 with $\psi(r,\omega) = 1(r \le \zeta_\beta^n)\phi(B_{r-}^\beta)$ and (b) of Lemma 3.4, the latter to check integrability, show that the following is a martingale (note that $B_r^\beta = B^\beta$ when $T_{\pi\beta} \le r < T_\beta$):

$$\begin{aligned} M_t &= \frac{1}{N(2N+\theta)} \sum_\beta 1(T_\beta \le t, T_\beta \le \zeta^n_\beta)\phi(B^\beta) \\ &\quad -\frac{1}{N}\int_0^t \sum_\beta 1(T_{\pi\beta} \le r < T_\beta, r \le \zeta^n_\beta)\phi(B^\beta)dr \\ &= \frac{1}{N(2N+\theta)} \sum_\beta 1(T_\beta \le t, T_\beta \le \zeta^n_\beta)\phi(B^\beta) - \int_0^t X^n_r(\phi)dr \end{aligned}$$

Using the definition of $[M]_t$ and reversing the last simplification we have

$$\begin{aligned} [M]_t &- \frac{1}{N(2N+\theta)}\int_0^t X^n_r(\phi^2)dr \\ &= \frac{1}{N^2(2N+\theta)^2} \sum_\beta 1(T_\beta \le t, T_\beta \le \zeta^n_\beta)\phi^2(B^\beta) \\ &\quad -\frac{1}{N(2N+\theta)}\int_0^t \frac{1}{N}\sum_\beta 1(T_{\pi\beta} \le r < T_\beta, r \le \zeta^n_\beta)\phi^2(B^\beta)dr \end{aligned}$$

is a martingale, whence $\langle M\rangle_t = \frac{1}{N(2N+\theta)}\int_0^t X^n_r(\phi^2)dr$.

Using the L^2 maximal inequality for martingales with the fact that $M_t^2 - \langle M\rangle_t$ is a martingale that is null at 0 gives

$$E\left(\sup_{s\le t} M_s^2\right) \le CE\langle M\rangle_t \tag{3.2}$$

Combining this with the formula for $\langle M\rangle_t$ and using Lemma 2.9, we get

$$E\left(\sup_{s\le t} M_s^2\right) \le CE\langle M\rangle_t \le \frac{C}{N(2N+\theta)}\|\phi\|_\infty^2 \int_0^t e^{\theta r} X^0_0(1)dr$$

Since $D^{n,1}_s(\phi) - \theta\int_0^s X^n_r(\phi)\,dr = \theta M_s$, the desired result follows. □

The third term from Section 2 that we will consider is:

$$D^{n,2}_t(\phi) = \frac{N+\theta}{N(2N+\theta)} \sum_\beta 1\left(T_\beta \le t, T_\beta \le \zeta^n_\beta\right)\left[\phi(B^\beta + W_\beta) - \phi(B^\beta)\right]$$

Lemma 2.3. *For all $t > 0$*

$$\lim_{N\to\infty} E\left(\sup_{s\le t}\left|D^{n,2}_s(\phi) - \int_0^s X^n_r(\Delta\phi/6)dr\right|\right) = 0$$

Proof. Given z and y in $\mathbf{R}^d$, we can apply the one-dimensional Taylor's theorem with remainder to $f(t) = \phi(y + t(z-y))$ to get

$$\phi(z)-\phi(y) = \sum_{i=1}^{d} \phi_i(y)(z_i-y_i)+\frac{1}{2}\sum_{1\le i,j\le d} \phi_{ij}(v)(z_i-y_i)(z_j-y_j) \quad (3.3)$$

where ϕ_i and ϕ_{ij} denote partial derivatives and v is a point on the line segment from y to z. Using this result, taking conditional expectation, and recalling that the vector W_β is independent of $\mathscr{F}_{T_\beta-}$ with $E(W^i_\beta) = 0$ and $EW^i_\beta W^j_\beta = 0$ for $i \neq j$ we have

$$E\big(\phi(B^\beta + W_\beta) - \phi(B^\beta) \,\big|\, \mathscr{F}_{T_\beta-}\big) = \frac{1}{2}\sum_{i=1}^{d} \phi_{ii}(B^\beta)E(W^i_\beta)^2 + R^\beta_N(\omega)$$

$$R^\beta_N(\omega) = \frac{1}{2}E\left(\sum_{1\le i,j\le d} \left[\phi_{ij}(v(\omega)) - \phi_{ij}(B^\beta)\right]^2 W^i_\beta W^j_\beta \,\middle|\, \mathscr{F}_{T_\beta-}\right)$$

Since $|v_i(\omega) - B^\beta_i| \le N^{-1/2}$ and $|W^i_\beta| \le N^{-1/2}$ for each i, and $\phi \in C^3_b$ it follows that

$$|R^\beta_N(\omega)| \le C(N^{-1/2})^2 \cdot (N^{-1/2})^2 \quad (3.4)$$

Now as $N \to \infty$, $\sqrt{N}W^i_\beta$ converges to a uniform distribution on $[-1, 1]$ which has second moment 1/3, so if we let $\nabla_\beta\phi = \phi(B^\beta + W_\beta) - \phi(B^\beta)$ then

$$E\big(\nabla_\beta\phi \,|\mathscr{F}_{T_\beta-}\big) = \left(\frac{1}{6} + \eta_N\right) N^{-1}\Delta\phi(B^\beta) + R^\beta_N(\omega) \quad (3.5)$$

where $\eta_N \to 0$ as $N \to \infty$. Applying Lemma 3.5 with

$$G_\beta = 1(T_\beta \le \zeta^n_\beta)\left\{\nabla_\beta\phi - E(\nabla_\beta\phi|\mathscr{F}_{T_\beta-})\right\}$$

and $K = cN^{-1/2}$ (ϕ is certainly Lipschitz continuous) we see that

$$M_t = N^{-1}\sum_\beta 1(T_\beta \le t)G_\beta$$

is a martingale with

$$E\Big(\sup_{s\le t} M^2_s\Big) \le \frac{C}{N}X^0_0(1)$$

Using the martingale M_t with (3.5) we can write

$$D^{n,2}_t(\phi) = \frac{N+\theta}{2N+\theta}\cdot M_t + D^{n,3}_t(\phi) + E^{n,2}_t(\phi)$$

$$D^{n,3}_t(\phi) = \frac{N+\theta}{2N+\theta}N^{-1}\left(\frac{1}{6} + \eta_N\right)\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta^n_\beta)\frac{1}{N}\Delta\phi(B^\beta)$$

$$E^{n,2}_t(\phi) = \frac{N+\theta}{2N+\theta}N^{-1}\sum_\beta 1(T_\beta \le t, T_\beta \le \zeta^n_\beta)R^\beta_N(\omega)$$

To handle $E_t^{n,2}(\phi)$ we observe that (3.4) and (b) in Lemma 3.4 imply

$$E\left(\sup_{s\le t}|E_s^{n,2}(\phi)|\right) \le \frac{C}{N} r(\theta,t) X_0^0(1) \to 0$$

as $N \to \infty$. For $D_t^{n,3}(\phi)$, we note that Lemma 2.2 (which only requires the boundedness of $\Delta\phi$) implies

$$E\left(\sup_{s\le t}\left|D_s^{n,3}(\phi) - \frac{N+\theta}{N}\left(\frac{1}{6}+\eta_N\right)\int_0^t X_r^n(\Delta\phi/6)dr\right|\right) \to 0$$

The result is now an easy consequence of the above estimates and Lemma 2.9. □

We turn our attention now to the fourth and final term:

$$E_t^{n,1}(\phi)$$
$$= \frac{1}{N}\sum_\beta a_\beta^n(t) h_\beta \left[\phi(B^\beta + W_\beta) 1\{B^\beta + W_\beta \notin \mathrm{supp}(X_{T_\beta -}^{n-1})\} - \phi(B^\beta)\right]$$

where $a_\beta^n(t) = 1(T_\beta \le t, T_{\pi\beta} < \zeta_\beta^n)$ and $h_\beta = 1(\delta_\beta = 1) - (N+\theta)/(2N+\theta)$.

Lemma 2.4. *For all* $t > 0$ $\lim_{N\to\infty} E\left(\sup_{s\le t}|E_s^{n,1}(\phi)|\right) = 0$

Proof. Lemma 3.5 implies that $E_t^{n,1}(\phi)$ is a martingale. To apply Lemma 3.5 here, first condition the β summand with respect to $\mathscr{F}_{T_\beta -} \vee \sigma(W_\beta)$. Using the L^2 maximal inequality, (3.1), and the trivial comparison $(a+b)^2 \le 2a^2 + 2b^2$, we have

$$\begin{aligned} E\left(\sup_{s\le t}|E_t^{n,1}(\phi)|^2\right) &\le CE([E^{n,1}(\phi)]_t) \\ &\le CE\Bigg(N^{-2}\sum_\beta a_\beta^n(t)\Big[\left\{\phi(B^\beta + W_\beta) - \phi(B^\beta)\right\}^2 \\ &\qquad + \phi^2(B^\beta) 1\left\{B^\beta + W_\beta \in \mathrm{supp}(X_{T_\beta -}^{n-1})\right\}\Big]\Bigg) \end{aligned} \tag{3.6}$$

Since any $\phi \in C_b^3$ is Lipschitz continuous and $|W_\beta^i| \le N^{-1/2}$

$$\{\phi(B^\beta + W_\beta) - \phi(B^\beta)\}^2 \le C|W_\beta|^2 \le \frac{C}{N}$$

Using (b) of Lemma 3.4, and $a_\beta^n(t) \le a_\beta^0(t)$, it follows that the first term in (3.6) is bounded by

$$E\left(N^{-3}\sum_{\beta} a^0_{\beta}(t)\right) \le N^{-1}Cr(\theta,t)X^0_0(1) \to 0 \quad \text{as} \quad N \to \infty$$

The second term in (3.6) is more complicated because we have to show it is small by showing that the indicator function is small, i.e., it is 0 most of the time. The proof of Lemma 2.4 will be completed once we establish Lemma 2.5. Section 4 is devoted to that task.

4. Upper bounds for the collision term

This section is devoted to the proof of Lemma 2.5. To begin we note that when a collision occurs, some particle β gave birth at the time of its death, T_β, onto a site occupied by at least one other particle γ who must have been born earlier and is still alive. In symbols, the expected value in Lemma 2.5 can be bounded above by

$$E\left(N^{-1}\sum_{\beta,\gamma} 1\left\{T_\beta \le t,\, T_{\pi\gamma} < T_\beta \le T_\gamma,\, B^\beta + W_\beta = B^\gamma \ne \Delta\right\}\right) \quad (4.1)$$

Define the σ-field of all events in the family line of α strictly before T_α, plus the value of t_α by

$$\mathscr{H}_\alpha = \sigma(t_{\alpha|m}, \delta_{\alpha|m}, e_{\alpha|m}, W_{\alpha|m} : m < |\alpha|) \vee \sigma(t_\alpha)$$

Conditioning on $\mathscr{H}_{\beta,\gamma} = \mathscr{H}_\beta \vee \mathscr{H}_\gamma$, we can rewrite (4.1) as

$$E\Big(N^{-1}\sum_{\beta,\gamma} 1\{T_\beta \le t,\, T_{\pi\gamma} < T_\beta \le T_\gamma\}\frac{1}{\psi(N)}1\{B^\gamma - B^\beta \in \mathscr{N}_N\}\Big) \quad (4.2)$$

where $\mathscr{N}_N = [-N^{-1/2}, N^{-1/2}]^d \cap \mathscr{Z}_N - \{0\}$ is the set of neighbors of 0, and $\psi(N) = |\mathscr{N}_N|$ is the number of neighbors. Here note that the inequality $T_{\pi\gamma} < T_\beta \le T_\gamma$ implies that γ is not a strict descendant of β and so on this $\mathscr{H}_{\beta,\gamma}$-measurable set we have

$$P(W_\beta \in \cdot|\mathscr{H}_{\beta,\gamma}) = P(W_\beta \in \cdot)$$

The reader should note that $B^\gamma - B^\beta \in \mathscr{N}_N$ implies that in particular that $B^\beta \ne \Delta$, $B^\gamma \ne \Delta$, i.e., β and γ are alive in the branching random walk.

Our final bit of notation before getting down to the work of doing the estimates is to let

$$\text{nbr}_{\beta,\gamma}(r) = 1(T_{\pi\beta} < r \le T_\beta, T_{\pi\gamma} < r \le T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N)$$

be the indicator of the event that β and γ are alive at time r in the branching random walk X^0 and they are neighbors. With this notation, we can use Lemma 3.2 with $\psi(r) = \text{nbr}_{\beta,\gamma}(r)$ (and with (a) of Lemma 3.4 to justify integrability) to rewrite (4.2) as

$$\begin{aligned} &N^{-1}\int_0^t (2N+\theta)\sum_{\beta,\gamma}\frac{1}{\psi(N)}E[\text{nbr}_{\beta,\gamma}(r)]dr \\ &\quad\le \frac{C}{\psi(N)}\int_0^t \sum_{\beta,\gamma} E[\text{nbr}_{\beta,\gamma}(r)]dr \end{aligned} \tag{4.3}$$

To estimate $E[\text{nbr}_{\beta\gamma}(r)]$ we need to consider the time and location of the most recent common ancestor of β and γ. The simplest situation is when $\beta_0 \ne \gamma_0$.

Lemma 4.1. *There is a constant* $0 < C < \infty$ *so that for all* $r \ge 0$

$$E\left[\sum_{\beta,\gamma:\gamma_0\ne\beta_0} \text{nbr}_{\beta,\gamma}(r)\right] \le Ce^{2\theta r}\cdot\{NX_0^0(1)\}^2\cdot[1+4(N+\theta)r]^{-d/2}$$

Before tackling the proof of this result we need some preliminaries. Let V_n^N be the random walk that with probability 1/2 stays put, and with probability 1/2 takes a jump uniformly distributed on $N^{1/2}\mathcal{N}_N$. We multiply by $N^{1/2}$ here so that as $N \to \infty$, V_m^N converges to V_m, that with probability 1/2 stays put, and with probability 1/2 takes a jump uniformly distributed on $[-1,1]^d$. The local central limit theorem for V_m implies that

$$P(V_m \in [-1,1]^d) \sim Cm^{-d/2} \quad \text{as} \quad m \to \infty$$

The next result which is (4) in Section 2 of Bramson, Durrett, and Swindle (1989), gives an upper bound that is uniform in N. It comes from a concentration function inequality of Kesten (1969).

Lemma 4.2. *There is a constant C independent of N so that if* $m \ge 0$ *then*

$$P(V_m^N \in x + [-1,1]^d) \le C(1+m)^{-d/2}$$

To convert the discrete time estimate in Lemma 4.2 to continuous time, we will use the following easily proved fact about the Poisson distribution.

Lemma 4.3. *Let $p > 0$. There is a constant $0 < C_p < \infty$ so that if $\lambda \geq 0$ then*

$$\sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!} (1+m)^{-p} \leq C_p (1+\lambda)^{-p}$$

Proof of Lemma 4.3. The trivial inequality $(1+m)^{-p} \leq 1$ and standard large deviations results for the Poisson distribution (see e.g., Ex. 1.4 on page 82 of Durrett (1995a)) imply that

$$\sum_{m=0}^{\lambda/2} e^{-\lambda} \frac{\lambda^m}{m!} (1+m)^{-p} \leq \sum_{m=0}^{\lambda/2} e^{-\lambda} \frac{\lambda^m}{m!} \leq e^{-c\lambda}$$

for some $c > 0$. The desired result follows since we have

$$\sum_{m>\lambda/2} e^{-\lambda} \frac{\lambda^m}{m!} (1+m)^{-p} \leq \left(1+\frac{\lambda}{2}\right)^{-p} \sum_{m>\lambda/2} e^{-\lambda} \frac{\lambda^m}{m!} \leq \left(1+\frac{\lambda}{2}\right)^{-p} \qquad \square$$

Proof of Lemma 4.1. When $i \neq j$,

$$\begin{aligned} &\sum_{\beta:\beta_0=i} \sum_{\gamma:\gamma_0=j} E[\mathrm{nbr}_{\beta,\gamma}(r)] \\ &\quad \leq \sum_{\beta:\beta_0=i} \sum_{\gamma:\gamma_0=j} P(T_{\pi\beta} < r < T_\beta, B^\beta \neq \Delta, T_{\pi\gamma} < r < T_\gamma, B^\gamma \neq \Delta) \\ &\qquad\qquad \times P\Big(N^{1/2}(x_j - x_i) + V^N_{|\beta|+|\gamma|} \in [-1,1]^d\Big) \end{aligned} \tag{4.4}$$

Breaking things down according to the values of $|\beta| = \ell$ and $|\gamma| = m$, we can write the last sum as

$$\begin{aligned} &\sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} \left(\frac{N+\theta}{2N+\theta}\right)^{\ell+m} \cdot 2^\ell \cdot 2^m \cdot e^{-2(2N+\theta)r} \frac{((2N+\theta)r)^{\ell+m}}{\ell!\, m!} \\ &\quad \cdot P(N^{1/2}(x_j - x_i) + V^N_{\ell+m} \in [-1,1]^d) \end{aligned} \tag{4.5}$$

The first factor gives the probability of no death along each line. The second and third the number of choices for β and γ. The fourth gives the probability that both β and γ are alive at time r.

Changing variables to $n = \ell + m$ and m and recalling $\sum_{m=0}^{n} n!/(n-m)!m! = 2^n$, we may convert (4.5) into

$$\sum_{n=0}^{\infty} \left(\frac{2N+2\theta}{2N+\theta}\right)^n \cdot e^{-2(2N+\theta)r} \frac{(2(2N+\theta)r)^n}{n!}$$

$$\cdot P(N^{1/2}(x_j - x_i) + V_n^N \in [-1,1]^d) \tag{4.6}$$

Doing some arithmetic and then using the upper bound in Lemma 4.2 we see this is no more than

$$\begin{aligned} &Ce^{2\theta r} \sum_{n=0}^{\infty} e^{-2(2N+2\theta)r} \frac{(2(2N+2\theta)r)^n}{n!} (1+n)^{-d/2} \\ &\le Ce^{2\theta r}(1+4(N+\theta)r)^{-d/2} \end{aligned}$$

by Lemma 4.3. This bound holds for each pair of values i and j. Multiplying by the square of the number of initial particles, $\{NX_0^0(1)\}^2$, gives the desired conclusion. □

To state the bound for the more difficult case in which $\gamma_0 = \beta_0$, we need to define

$$I(u) = 1 + \int_0^u (1+x)^{-d/2}\, dx$$

This will often be compared with $\psi_0(N) = \psi(N)/N$, so we note now that the definition of $\psi(N)$ and a little calculus show that there are constants $0 < c < C < \infty$ so that

$$c\psi_0(N) \le I(N) \le C\psi_0(N) \quad \text{for} \quad N \ge 1 \tag{4.7}$$

If $\gamma_0 = \beta_0$, $\beta \wedge \gamma$ denotes the most recent common ancestor of β and γ, i.e., the unique ancestor of γ and β which maximizes $|\beta \wedge \gamma|$, and if $\beta_0 \ne \gamma_0$ set $\beta \wedge \gamma = \emptyset$.

Lemma 4.4. *There is a constant* $0 < C < \infty$ *so that*

$$E\Big[\sum_{\beta,\gamma} 1(\beta_0 = \gamma_0)\, \mathrm{nbr}_{\beta,\gamma}(r)\Big] \le Ce^{2\theta r}\{NX_0^0(1)\} \cdot I(2(N+\theta)r)$$

Before we get involved in the details of the proof of Lemma 4.2, we will do the

Proof of Lemma 2.5. Using Lemmas 4.1 and 4.4 with (4.3) we may bound the expectaion in Lemma 2.5 by

$$\begin{aligned} &C\psi(N)^{-1} \cdot \{NX_0^0(1)\}^2 \int_0^t \frac{e^{2\theta r}}{(1+4(N+\theta)r)^{d/2}}\, dr \\ &\quad + C\,\psi(N)^{-1} \cdot \{NX_0^0(1)\} \int_0^t e^{2\theta r} I(2(N+\theta)r) dr \end{aligned}$$

Changing variables $s = 1 + (4N + 4\theta)r$, $dr = ds/(4N + 4\theta)$ in the first integral and using a trivial bound on the second which has an increasing integrand we see the above equals

$$Ce^{2\theta t} \cdot \{X_0^0(1)\}^2 \cdot \frac{1}{\psi_0(N)} \cdot \int_1^{1+(4N+4\theta)t} s^{-d/2}\, ds$$
$$+ Cte^{2\theta t} \cdot \{X_0^0(1)\} \cdot \frac{1}{\psi_0(N)} \cdot I(2(N+\theta)t)$$

The integral is bounded by $I((4N + 4\theta)t)$, so the desired result follows from (4.7) (recall our convention about constants C). □

Proof of Lemma 4.4. Note that $\beta_0 = \gamma_0$, $B^\beta - B^\gamma \neq 0$, $T_{\pi\beta} < r \leq T_\beta$, and $T_{\pi\gamma} < r \leq T_\gamma$ imply $\gamma \wedge \beta$ is not $\emptyset$, β, or γ. Let $k < |\beta| \wedge |\gamma|$ be such that $\beta \wedge \gamma = \beta|k = \gamma|k$. Let $\ell \geq 1$ be such that $|\gamma| = k + \ell$. On $\{T_{\pi\beta} < r \leq T_\beta, B^\beta \neq \Delta\}$ we have

$$E\left(\sum_{\substack{\gamma:|\beta\wedge\gamma|=k,\\ |\gamma|=k+\ell}} 1\{T_{\pi\gamma} < r \leq T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\} | \mathcal{H}_\beta \right)$$
$$\leq e^{-(2N+\theta)(r-T_{\beta|k})} \frac{((2N+\theta)(r - T_{\beta|k}))^{\ell-1}}{(\ell-1)!}$$
$$\cdot \left(\frac{N+\theta}{2N+\theta}\right)^{\ell-1} \cdot 2^{\ell-1} \cdot C((\ell-1)+1)^{-d/2}$$

The first factor on the right-hand side reflects the fact that there must be exactly $\ell - 1$ more generations in the γ line at time r. The second that there can be no deaths along the way. The third, $2^{\ell-1}$, gives the number of γ with the properties stated in the sum. The fourth comes from Lemma 4.2. Summing the last result over $\ell \geq 1$ gives

$$E\left(\sum_{\substack{\gamma:|\beta\wedge\gamma|=k,\\ |\gamma|>k}} 1\{T_{\pi\gamma} < r \leq T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\} | \mathcal{H}_\beta \right)$$
$$\leq Ce^{-(2N+\theta)(r-T_{\beta|k})} \sum_{\ell=1}^{\infty} \frac{[2(N+\theta)(r-T_{\beta|k})]^{\ell-1}}{(\ell-1)!} (1 + (\ell-1))^{-d/2} \tag{4.8}$$

Using Lemma 4.3 now, it follows that (4.8) is at most

$$Ce^{\theta r}\left[1 + 2(N+\theta)(r - T_{\beta|k})\right]^{-d/2}$$

Summing over $0 \le k < |\beta|$ we have that on $\{T_{\pi\beta} < r \le T_\beta, B^\beta \ne \Delta\}$ that

$$E\left(\sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < r \le T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\}|\mathcal{H}_\beta\right) \le Ce^{\theta r} H(\beta, r) \tag{4.9}$$

where $H(\beta, r) = \sum_{k=0}^{|\beta|-1}[1 + 2(N+\theta)(r - T_{\beta|k})]^{-d/2}$. Since $r > T_{\pi\beta}$ we have

$$H(\beta, r) \le \sum_{k=0}^{|\beta|-1}[1 + 2(N+\theta)(T_{\pi\beta} - T_{\beta|k})]^{-d/2} \equiv H(\beta)$$

Summing over β we see that

$$E\left(\sum_{\beta,\gamma} 1(\beta_0 = \gamma_0)\mathrm{nbr}_{\beta,\gamma}(r)\right) \le Ce^{\theta r} E\left(\sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, B^\beta \ne \Delta\}H(\beta)\right) \tag{4.10}$$

To estimate (4.10) we will break the sum down according to the value of $|\beta| = m$. (Note that by the remarks at the beginning of the proof we must have $m \ge 1$.) Let $\xi_1, \xi_2, \ldots$ be independent mean 1 exponentials, and let $\Gamma_m = \xi_1 + \cdots + \xi_m$. To explain our choice of notation, observe that Γ_m has a gamma$(m, 1)$ distribution. Using our new symbols, we can bound the right-hand side of (4.10) (through a now familiar argument) by

$$Ce^{\theta r}\{NX_0^0(1)\}\sum_{m=1}^\infty \left(\frac{2(N+\theta)}{2N+\theta}\right)^m e_m((2N+\theta)r) \tag{4.11}$$

where

$$e_m(u) = E\left(1\{\Gamma_m < u < \Gamma_{m+1}\}\sum_{k=0}^{m-1}[1 + \Gamma_m - \Gamma_{k+1}]^{-d/2}\right) \tag{4.12}$$

To check the indexing of the Γ's here, note that for $k \ge 0$, $T_{\beta|k}$ is the time of the $(k+1)$th death along the line of descent of β.

Evaluating (4.12) and then the sum in (4.11) is a simple (though somewhat tedious) exercise about the rate one Poisson process. To make the result available for later use, we recall $\epsilon_N = \theta/(2N+\theta)$ and the function I defined prior to Lemma 4.4, and state

Lemma 4.5. *For all* $u > 0$ $\quad \sum_{m=1}^{\infty}(1+\epsilon_N)^m e_m(u) \le 3e^{u\epsilon_N} I(u)$

Setting $u = (2N+\theta)r$ in Lemma 4.5, and using (4.10)–(4.12) gives Lemma 4.4. Thus we can complete its proof by doing the

Proof of Lemma 4.5. We begin by biting off a small part of the problem. The reason for doing this will become clear when we tackle the main piece. When $k = m-1$, $\Gamma_m = \Gamma_{k+1}$ and the term in the sum is $1^{-d/2} = 1$. Let

$$f_m(u) = P(\Gamma_m < u < \Gamma_{m+1}) \quad \text{and} \quad \hat{e}_m(u) = e_m(u) - f_m(u)$$

Summing this contribution of the $k = m-1$ term for $m \ge 1$ we have

$$\begin{aligned} &\sum_{m=1}^{\infty} (1+\epsilon_N)^m P(\Gamma_m < u < \Gamma_{m+1}) \\ &\quad = \sum_{m=1}^{\infty} (1+\epsilon_N)^m e^{-u} \frac{u^m}{m!} \le \exp(u\epsilon_N) \end{aligned} \tag{4.13}$$

To begin to tackle the main piece we note that

$$\sum_{k=0}^{m-2} \frac{x^{m-2-k}}{(m-2-k)!} \frac{y^k}{k!} = \frac{(x+y)^{m-2}}{(m-2)!} \tag{4.14}$$

Summing over $m \ge 2$ now we have

$$\sum_{m=2}^{\infty}(1+\epsilon_N)^m \cdot \sum_{k=0}^{m-2} \frac{x^{m-2-k}}{(m-2-k)!} \frac{y^k}{k!} = (1+\epsilon_N)^2 e^{(x+y)(1+\epsilon_N)} \tag{4.15}$$

Using the fact that $\Gamma_m - \Gamma_{k+1} (= x)$ and $\Gamma_{k+1} (= y)$ have independent gamma distributions, one can write

$$\begin{aligned} &E\left(1\{\Gamma_m < u < \Gamma_{m+1}\} \sum_{k=0}^{m-2} [1 + \Gamma_m - \Gamma_{k+1}]^{-d/2}\right) \\ &\quad = \sum_{k=0}^{m-2} \int_0^{\infty} \int_0^{\infty} 1\{x+y \le u\} e^{-u+x+y} (1+x)^{-d/2} \\ &\qquad\qquad \times e^{-x} \frac{x^{m-k-2}}{(m-k-2)!} \cdot e^{-y} \frac{y^k}{k!}\, dx\, dy \end{aligned} \tag{4.16}$$

Summing the last estimate and using (4.15) gives

$$\sum_{m=1}^{\infty}(1+\epsilon_N)^m \hat{e}_m(u) \le \int_0^u dx \int_0^{u-x} dy\, e^{-u} (1+x)^{-d/2} e^{(x+y)(1+\epsilon_N)} (1+\epsilon_N)^2 \tag{4.17}$$

The inside integral is $e^{-u}\int_0^{u-x} dy\, e^{(x+y)(1+\epsilon_N)}(1+\epsilon_N)^2 \le e^{u\epsilon_N}(1+\epsilon_N) \le 2e^{u\epsilon_N}$, so the right hand side of (4.17) is bounded by

$$2e^{u\epsilon_N}\int_0^u (1+x)^{-d/2}\,dx \tag{4.18}$$

Adding (4.18) to (4.13) gives the desired result. □

Later to estimate second moments of the collision term, we will need an estimate for

$$g_m(u) = E\left\{1\{\Gamma_m < u < \Gamma_{m+1}\}\left(\sum_{k=0}^{m-1}[1+\Gamma_m-\Gamma_{k+1}]^{-d/2}\right)^2\right\} \tag{4.19}$$

The methods are similar to the proof just completed, so we will give the proof here.

Lemma 4.6. There is a C such that for all $u > 0$

$$\sum_{m=1}^{\infty}(1+\epsilon_N)^m g_m(u) \le Ce^{2\epsilon_N^+ u} I(u)^2$$

Proof. We may assume $\theta \ge 0$ (so that $\epsilon_N \ge 0$) because the result for $\theta < 0$ clearly follows from the result for $\theta = 0$. Reversing the order of the first m increments we can be write

$$g_m(u) = E\left\{1\{\Gamma_m < u < \Gamma_{m+1}\}\left(\sum_{j=0}^{m-1}(1+\Gamma_j)^{-d/2}\right)^2\right\}$$

Writing the square as a double sum and counting the diagonal twice, the above is at most

$$2E\left(\sum_{0\le j\le k<m} 1\{\Gamma_m < u < \Gamma_{m+1}\}(1+\Gamma_j)^{-d/2}(1-\Gamma_k)^{-d/2}\right)$$

Multiplying by $(1+\epsilon_N)^m$, summing over m and doing some rearrangement gives

$$\sum_{m=1}^{\infty}(1+\epsilon_N)^m g_m(u) \le 2E\left(\sum_{j=0}^{\infty}(1+\Gamma_j)^{-d/2}\sum_{k=j}^{\infty}(1+\Gamma_k)^{-d/2}\right.$$
$$\left.\times\sum_{m>k}(1+\epsilon_N)^m 1\{\Gamma_m < u < \Gamma_{m+1}\}\right) \tag{4.20}$$

To bound this we begin with the inner sum. On $\{\Gamma_k < u\}$ we can change variables $i = m - k$ and $\Gamma'_i = \Gamma_{k+i} - \Gamma_k$ to get

$$E\left(\sum_{m>k} (1+\epsilon_N)^m 1\{\Gamma_m < u < \Gamma_{m+1}\} \middle| \Gamma_1, \dots \Gamma_k \right)$$
$$= (1+\epsilon_N)^k E\left(\sum_{i\ge 1} (1+\epsilon_N)^i 1\{\Gamma'_i < u - \Gamma_k < \Gamma'_{i+1}\} \middle| \Gamma_1, \dots \Gamma_k \right)$$

Since Γ'_i is independent of $\Gamma_1, \dots \Gamma_k$, the above equals

$$(1+\epsilon_N)^k \sum_{i=1}^{\infty} (1+\epsilon_N)^i e^{-(u-\Gamma_k)} \frac{(u-\Gamma_k)^i}{i!} \le (1+\epsilon_N)^k e^{\epsilon_N (u-\Gamma_k)}$$
$$\le (1+\epsilon_N)^k e^{u\epsilon_N} \quad ,$$

where we have used the hypothesis $\epsilon_N \ge 0$ in the last inequality. Using this in (4.20) and isolating the $k = j$ and $j = 0$ terms we have

$$\sum_{m=1}^{\infty} (1+\epsilon_N)^m g_m(u) \le 2e^{u\epsilon_N} \left[\sum_{j=0}^{\infty} (1+\epsilon_N)^j E\left\{1(\Gamma_j < u)(1+\Gamma_j)^{-d}\right\} \right.$$
$$+ \sum_{k=0}^{\infty} (1+\epsilon_N)^k E\left\{1(\Gamma_k < u)(1+\Gamma_k)^{-d/2}\right\}$$
$$+ \sum_{j=1}^{\infty} \sum_{k=j+1}^{\infty} (1+\epsilon_N)^k$$
$$\left. \times E\{1(\Gamma_k < u)(1+\Gamma_j)^{-d/2}(1+\Gamma_k)^{-d/2}\} \right] \tag{4.21}$$

Since $1 + \Gamma_j \ge 1$, the first sum on the right is smaller than the second. To bound the second sum in (4.21) we note that

$$\sum_{k=0}^{\infty} (1+\epsilon_N)^k E\left\{1(\Gamma_k < u)(1+\Gamma_k)^{-d/2}\right\}$$
$$= 1 + \sum_{k=1}^{\infty} (1+\epsilon_N)^k \int_0^u e^{-x} \frac{x^{k-1}}{(k-1)!} (1+x)^{-d/2}\, dx$$
$$\le 1 + (1+\epsilon_N) e^{u\epsilon_N} \int_0^u (1+x)^{-d/2}\, dx \quad , \tag{4.22}$$

again using $\epsilon_N \ge 0$ in the last inequality. Using the fact that $\Gamma_j (= x)$ and $\Gamma_k - \Gamma_j (= y)$ have independent gamma distributions, we see that the double sum in (4.21) is

$$\int_0^\infty dx \int_0^\infty dy\, 1(x+y \le u)(1+x)^{-d/2}(1+x+y)^{-d/2} \times \sum_{0<j<k} e^{-x-y} \frac{x^{j-1}}{(j-1)!} \frac{y^{k-j-1}}{(k-j-1)!}(1+\epsilon_N)^k \tag{4.23}$$

Using analogues of (4.14) and (4.15) now with the fact that $x + y \le u$, we see that the double sum in (4.23) is smaller than $(1 + \epsilon_N)e^{u\epsilon_N}$. Replacing $(1 + x + y)^{-d/2}$ by $(1 + y)^{-d/2}$ and enlarging the domain of integration, (4.23) is bounded by

$$(1+\epsilon_N)e^{u\epsilon_N} \int_0^u dx \int_0^u dy\, (1+x)^{-d/2}(1+y)^{-d/2} \le 2e^{u\epsilon_N} I(u)^2 \tag{4.24}$$

Combining (4.21)–(4.24) gives Lemma 4.6. □

5. Collisions between distant relatives can be ignored

In this section we will refine the bounds in Lemmas 4.1 and 4.4 proving the claim in the section's name. Imitating (4.2), we can bound the collision term for unrelated individuals by

$$J_0(t) = \frac{1}{N\psi(N)} \sum_{\beta,\gamma:\beta_0 \ne \gamma_0} 1\{T_\beta \le t, T_{\pi\beta} < T_\beta < T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\} \tag{5.1}$$

If we start with all the particles in one neighborhood then this term will not be small. Thus to have $EJ_0(t) \to 0$, we must assume that the particles are sufficiently spread out in the initial distribution. The next result gives a simple sufficient condition.

Lemma 5.1. *If $X_0^{0,N}$ converges to X_0 in $M_F(\mathbf{R}^d)$ where X_0 is an atomless measure, then for any $t \ge 0$, $E(J_0(t)) \to 0$ as $N \to \infty$.*

Proof. As we converted (4.2) into (4.3), we can bound $E(J_0(t))$ by

$$\frac{C}{\psi(N)} \int_0^t \sum_{\beta,\gamma:\beta_0 \ne \gamma_0} E[\mathrm{nbr}_{\beta,\gamma}(r)]dr$$

The bound on this that results from Lemma 4.1 is

$$Ce^{2\theta t}\frac{\{NX_0^{0,N}(1)\}^2}{\psi(N)}\int_0^t [1+4(N+\theta)r]^{-d/2}\,dr \tag{5.2}$$

This is almost good enough by itself, but clearly we need to get a better estimate for small (i.e., $O(1/N)$) values of r for which we will use the atomless assumption below.

Repeating the computations in (4.4)–(4.6) we see that if $i \neq j$

$$\begin{aligned}\sum_{\beta:\beta_0=i}\sum_{\gamma:\gamma_0=j} & E[\mathrm{nbr}_{\beta,\gamma}(r)] \\ & \leq \sum_{n=0}^{\infty}\left(\frac{2(N+\theta)}{2N+\theta}\right)^n \cdot e^{-2(2N+\theta)r}\frac{(2(2N+\theta)r)^n}{n!} \\ & \cdot P(N^{1/2}(x_j-x_i)+V_n^N \in [-1,1]^d)\end{aligned} \tag{5.3}$$

Using Lemma 4.2 to bound the last probability, then integrating over $[0,t]$, and summing over $i \neq j$, we see that the contribution to $J_0(t)$ from $n \geq N$ is at most

$$\begin{aligned}\{NX_0^{0,N}(1)\}^2 & \frac{1}{\psi(N)}\int_0^t \sum_{n=N}^{\infty} e^{-2(2N+\theta)r}\frac{(2(2N+2\theta)r)^n}{n!}\cdot CN^{-d/2}\,dr \\ & \leq \{NX_0^{0,N}(1)\}^2\frac{CN^{-d/2}}{\psi(N)}\int_0^t e^{2\theta r}\,dr\end{aligned} \tag{5.4}$$

The last integral is bounded by C. So considering the two possibilities: $d>2$ in which case $\psi(N)\sim CN$, and $d=2$ in which case $\psi(N)\sim CN\log N$, we see that (5.4) $\to 0$ as $N\to\infty$.

The contribution to $J_0(t)$ from $n \leq N$ can be bounded by

$$\begin{aligned}\frac{1}{\psi(N)}\int_0^t C\sum_{n=0}^{N} & e^{-2(2N+\theta)r}\,\frac{(2(2N+\theta)r)^n}{n!}\,dr \\ & \cdot\sum_{i\neq j} P\big(N^{1/2}(x_j-x_i)+V_n^N \in [-1,1]^d\big)\end{aligned}$$

Changing variables $s = 2(2N+\theta)r$, $dr = ds/2(2N+\theta)$ in the integral, and then noticing the gamma density $e^{-s}s^n/n!$ has total mass 1, we see the above is no more than

$$\frac{1}{\psi(N)}\cdot\frac{C}{2N+\theta}\sum_{n=0}^{N}\sum_{i\neq j}P\big(N^{1/2}(x_j-x_i)+V_n^N \in [-1,1]^d\big) \tag{5.5}$$

This quantity is easy to estimate in $d > 2$. Using the fact that $\psi(N) \sim CN$, and then using Lemma 4.2, we have that (5.5) is less than or equal to

$$\left(X_0^{0,N} \times X_0^{0,N}\right)(\{(x, y) : \|x - y\| \le \epsilon\}) \sum_{n=0}^{N} C(1+n)^{-d/2} + \left\{X_0^{0,N}(1)\right\}^2 \sum_{n=\epsilon N^{1/2}-1}^{N} C(1+n)^{-d/2} \tag{5.6}$$

where in the second term we have used the fact that if $\|x_i - x_j\| > \epsilon$ it takes at least $\epsilon N^{1/2} - 1$ steps of V_n^N to get to $N^{1/2}(x_i - x_j) + [-1, 1]^d$. Since $\{(x, y) : \|x - y\| \le \epsilon\}$ is a closed set, the limsup of the first term as $N \to \infty$ is bounded by

$$C(X_0 \times X_0)(\{(x, y) : \|x - y\| \le \epsilon\})$$

Since we have supposed that X_0 has no atoms, the above expression is small if ϵ is. Now in $d > 2$ the second term in (5.6) tends to 0 for any $\epsilon > 0$, since the sum converges.

(5.5) offers more resistance in the borderline case $d = 2$. Using $\psi(N) \sim CN\log(N)$, and then Lemma 4.2, but decomposing things into three pieces now, we see that (5.5) is at most

$$\begin{aligned} &\frac{C}{\log N}\left(X_0^{0,N} \times X_0^{0,N}\right)(\{(x, y) : \|x - y\| \le \epsilon\}) \sum_{n=0}^{N}(1+n)^{-1} \\ &\quad + \frac{C}{\log N}\sum_{n=\epsilon N^{1/2}-1}^{N/(\log N)^3} N^{-2} \sum_{i \ne j : \|x_i - x_j\| > \epsilon} P\left(N^{1/2}(x_j - x_i) + V_n^N \in [-1, 1]^d\right) \\ &\quad + \frac{C}{\log N}\{X_0^{0,N}(1)\}^2 \sum_{n=N/(\log N)^3+1}^{N} (1+n)^{-1} \end{aligned} \tag{5.7}$$

Again the limsup of the first term as $N \to \infty$ is bounded by

$$C(X_0 \times X_0)(\{(x, y) : \|x - y\| \le \epsilon\})$$

which is small if ϵ is. Bounding the sum of $(1+n)^{-1}$ by the integral of x^{-1} we see that the third term is no more than

$$\frac{C}{\log N}\{X_0^{0,N}(1)\}^2 (3 \log\log N) \to 0$$

To handle the second term in (5.7) we will use a standard estimate for "small" large deviations of random walks.

Lemma 5.2. *There are constants* $0 < \delta_0, c, C < \infty$ *so that if* $0 < z/n < \delta_0$, *then*

$$P(\|V_n^N\| \geq z) \leq C\exp(-cz^2/n)$$

Proof. If $S_n = \eta_1 + \cdots + \eta_n$ where the η_i are i.i.d. real random variables with mean 0 and $|\eta_i| \leq 1$ a.s., and $\theta > 0$ then

$$P(S_n > z) \leq e^{-z\theta}(E\exp(\theta\eta_1))^n$$

A Taylor expansion shows that $E(\exp(\theta\eta_1)) \leq 1 + e^{|\theta|}\theta^2/2$. Using this fact and the inequality with $\theta = z/n$ leads to

$$P(S_n > z) \leq \exp\left(\frac{-z^2}{n}[1 - e^{\delta_0}/2]\right)$$

Apply this to each coordinate of V_n^N and their negatives to obtain the desired result. □

Turn now to the second term in (5.7). Let $z = (\log N)n^{1/2}$ and use the fact that $n \geq \epsilon N^{1/2} - 1$ in the second sum in (5.7) to see that

$$z = n(\log N)/n^{1/2} \leq n(\log N)/(\epsilon N^{1/2} - 1)^{1/2} \leq \delta_0 n$$

for large N, so the conditions of Lemma 5.2 hold. Plugging in the chosen value of z

$$P(\|V_n^N\| \geq (\log N)n^{1/2}) \leq C\exp(-c(\log N)^2) = CN^{-c\log N}$$

To convert this into the result we need for (5.7) note that $n \leq N/(\log N)^3$ implies

$$z = (\log N)n^{1/2} \leq N^{1/2}/(\log N)^{1/2} \leq \frac{\epsilon}{2}N^{1/2} \ ,$$

so it follows that if $\|x_i - x_j\| > \epsilon$ then

$$P\big(N^{1/2}(x_j - x_i) + V_n^N \in [-1, 1]^d\big) \leq CN^{-c\log N}$$

This is more than enough to send the second term in (5.7) to 0. This completes the estimation of (5.5) in the case $d = 2$ and we have established Lemma 5.1. □

We turn now to the more difficult task of estimating the probability of collision for lines with $\beta_0 = \gamma_0$. If we fix an amount of time τ then we can define the collisions of related particles more distantly related than τ by

$$J(t,\tau) = \frac{1}{N\psi(N)} \sum_{\beta,\gamma:\beta_0=\gamma_0} 1\{T_\beta \le t, T_{\pi\gamma} < T_\beta \le T_\gamma\} \cdot 1\{T_{\beta\wedge\gamma} \le T_\beta - \tau\} \cdot 1\{B^\gamma - B^\beta \in \mathcal{N}_N\} \tag{5.8}$$

When $t \le \tau$ it is impossible to satisfy all the conditions inside the sum, so $J(t,\tau) = 0$.

Lemma 5.3. *There is a constant* $0 < C < \infty$, *depending on* t *and* θ (*according to our usual convention*) *so that for all* $\tau \le t$

$$E(J(t,\tau)) \le \frac{C\{X_0^{0,N}(1)\}}{\psi_0(N)} \int_{(2N+\theta)\tau}^{(2N+\theta)t} dy\,(1+y)^{-d/2}$$

Proof. As we converted (4.2) into (4.3), we can bound $J(t,\tau)$ by

$$\frac{C}{\psi(N)} \int_0^t \sum_{\beta,\gamma:\beta_0=\gamma_0} E[\mathrm{nbr}_{\beta,\gamma}(r) 1\{T_{\beta\wedge\gamma} \le r-\tau\}]dr \tag{5.9}$$

Now $\beta_0 = \gamma_0$, $B^\gamma - B^\beta \ne 0$, $T_{\pi\beta} < r \le T_\beta$, and $T_{\pi\gamma} < r \le T_\gamma$ imply that $\beta\wedge\gamma$ is not $\emptyset$, β, or γ. Let $0 \le k < |\beta|\wedge|\gamma|$ be such that $\beta\wedge\gamma = \beta|k = \gamma|k$. By conditioning on $\mathcal{H}_\beta$ and using Lemma 4.2 we can bound (5.9) by

$$\frac{C}{\psi(N)} \int_0^t \sum_\beta \sum_{k=0}^{|\beta|-1} E\Bigg(1\{T_{\beta|k} < r-\tau, T_{\pi\beta} < r \le T_\beta, B^\beta \ne \Delta\} \cdot (|\beta|-k)^{-d/2} \cdot E\Bigg(\sum_{\gamma:\gamma\wedge\beta=\beta|k} 1(T_{\pi\gamma} < r \le T_\gamma, B^\gamma \ne \Delta) \Bigg| \mathcal{H}_\beta \Bigg)\Bigg) dr \tag{5.10}$$

The conditional expectation is just the expected number of children at time r of the particle $\gamma|(k+1)$, and so by Lemma 2.9 it is $e^{\theta r} \le e^{|\theta|t}$. Using this and then evaluating $P(B^\beta \ne \Delta)$, we bound (5.10) by

$$\frac{Ce^{|\theta|t}}{\psi(N)} \int_0^t \sum_\beta \left(\frac{N+\theta}{2N+\theta}\right)^{|\beta|} \times \sum_{k=0}^{|\beta|-1} (|\beta|-k)^{-d/2} P(T_{\beta|k} < r-\tau, T_{\pi\beta} < r \le T_\beta)dr \tag{5.11}$$

Letting $\ell = |\beta|$, using our standard Gamma random variables Γ_m, and recalling $\epsilon_N = \theta/(2N+\theta)$, (5.11) can be written as

$$\begin{aligned}\frac{Ce^{|\theta|t}}{\psi(N)} \cdot \{NX_0^{0,N}(1)\} \int_0^t \sum_{\ell=0}^{\infty} (1+\epsilon_N)^{\ell} \sum_{k=0}^{\ell-1} (\ell-k)^{-d/2} \\ \times P(\Gamma_{k+1} < (2N+\theta)(r-\tau), \Gamma_\ell < (2N+\theta)r < \Gamma_{\ell+1})\, dr \qquad (5.12)\end{aligned}$$

If $\ell = k+1$ the last probability is just

$$\begin{aligned}&P(\Gamma_{k+1} < (2N+\theta)(r-\tau) < (2N+\theta)r < \Gamma_{k+2}) \\ &= \int_0^{(2N+\theta)(r-\tau)} e^{-x}\frac{x^k}{k!} \cdot e^{x-(2N+\theta)r}\, dx\end{aligned}$$

If $\ell > k+1$ we have to integrate out the value of $\Gamma_\ell - \Gamma_{k+1} = y$ and the result is

$$\int_0^{(2N+\theta)(r-\tau)} dx e^{-x}\frac{x^k}{k!} \int_0^{(2N+\theta)r-x} dy e^{-y} \frac{y^{\ell-k-2}}{(\ell-k-2)!} e^{x+y-(2N+\theta)r}$$

Interchanging the order of summation, setting $j = \ell - k - 2$, which runs from -1 (for $\ell = k+1$) to ∞, and using the above expressions, we can write the double sum in (5.12) as

$$\begin{aligned}&e^{-(2N+\theta)r} \int_0^{(2N+\theta)(r-\tau)} dx \sum_{k=0}^{\infty} (1+\epsilon_N)^{k+1} \frac{x^k}{k!} \\ &\cdot \left[1 + \int_0^{(2N+\theta)r-x} dy \sum_{j=0}^{\infty} (1+\epsilon_N)^{j+1} \frac{y^j}{j!} (j+2)^{-d/2}\right] \qquad (5.13)\end{aligned}$$

Doing the sum over k, estimating the sum over j using Lemma 4.3, and then absorbing the extra $(1+\epsilon_N)^2$ into the C we bound the above by

$$\begin{aligned}&Ce^{-(2N+\theta)r} \int_0^{(2N+\theta)(r-\tau)} dx\, e^{x(1+\epsilon_N)} \\ &\cdot \left[1 + \int_0^{(2N+\theta)r-x} dy\, e^{y(1+\epsilon_N)} (1+y)^{-d/2}\right] \qquad (5.14)\end{aligned}$$

A little calculus (left to the reader) shows that

Lemma 5.4. *There is a constant* $0 < C < \infty$ *so that if* $\eta \geq -1/2$

$$1 + \int_0^z e^{y(1+\eta)} (1+y)^{-d/2}\, dy \leq Ce^{z(1+\eta)} (1+z)^{-d/2}$$

Using this with $z = (2N+\theta)r - x$, and $\eta = \epsilon_N$, we see that (5.14) is no more than

$$Ce^{-(2N+\theta)r} \int_0^{(2N+\theta)(r-\tau)} e^{x(1+\epsilon_N)} \times e^{(1+\epsilon_N)\{(2N+\theta)r-x\}}(1+\{(2N+\theta)r-x\})^{-d/2}\,dx$$

Changing variables $y = (2N+\theta)r - x$, we may bound (5.13) by

$$Ce^{\theta r} \int_{(2N+\theta)\tau}^{(2N+\theta)r} (1+y)^{-d/2}\,dy$$

Inserting the last result into (5.12), we have an upper bound

$$EJ(t,\tau) \le \frac{C}{\psi(N)} \cdot \{NX_0^{0,N}(1)\} \cdot t \cdot e^{2|\theta|t} \int_{(2N+\theta)\tau}^{(2N+\theta)t} (1+y)^{-d/2}\,dy$$

which easily converts into the bound given in Lemma 5.3. □

6. Convergence of the collision term

The goal of this section is to analyze the limiting behavior of the collision term

$$K_t^n(\phi) = \frac{N+\theta}{N(2N+\theta)} \sum_\beta a_\beta^n(t)\phi(B^\beta + W_\beta)1\{B^\beta + W_\beta \in \operatorname{supp}(X^{n-1}_{T_\beta -})\}$$

where $a_\beta^n(t) = 1(T_\beta \le t,\ \zeta_\beta^n > T_{\pi\beta})$ is 1 if the particle was once alive in X^n but died before time t. More specifically we will commence the proof of Lemma 2.6 which we now restate as Theorem 6.1. Recall from (2.12) that our test functions ϕ belong to C_b^3.

Theorem 6.1. *For* $n = 1, 2$ *and any* $0 \le t < \infty$

$$\lim_{N\to\infty} E\left(\sup_{s\le t}\left|K_s^n(\phi) - \int_0^s b_d X_r^1(\phi)\,dr\right|\right) = 0 \tag{6.1}$$

We really do mean X^1 in the above and not X^n. Of course Proposition 1 and the ordering of the X^n's (see (2.8)) show the difference is unimportant. To prove Theorem 6.1, we will slowly change $K_s^n(\phi)$ into the integral. We first outline the main steps in a sequence of Lemmas and then will provide the proofs in this and the next four sections. In the first step, we tidy up the expression a little replacing $\phi(B^\beta + W_\beta)$ by $\phi(B^\beta)$. We also make a more significant change by replacing the collision events themselves by their conditional probabilities given the information available just before the displacement occurred. Recall that

$$\mathscr{N}_N = [-N^{-1/2}, N^{-1/2}]^d \cap \mathscr{L}_N - \{0\}$$

denotes the neighbors of 0 in our lattice $\mathscr{L}_N$ and $\psi(N) = |\mathscr{N}_N|$ is the number of neighbors. Let

$$\nu_m(\beta) = |\{B^\gamma : T_{\pi\gamma} < T_\beta \le T_\gamma,\ B^\gamma - B^\beta \in \mathscr{N}_N,\ \zeta^m_\gamma > T_{\pi\gamma}\}| \quad (6.2)$$

be the number of neighbors of B^β occupied in X^m at time $T_\beta-$. Write $\gamma \le \beta$ if γ is an ancestor of β and use $\gamma < \beta$ if it is a strict ancestor. For future reference note that the conditions $T_{\pi\gamma} < T_\beta \le T_\gamma$ and $B^\gamma - B^\beta \ne 0$ eliminate $\gamma \le \beta$ or $\beta \le \gamma$.

We define our first modification of the collision term by

$$K^{n,1}_t(\phi) = \frac{1}{2N+\theta} \sum_\beta a^n_\beta(t)\, \phi(B^\beta)\, \frac{\nu_{n-1}(\beta)}{\psi(N)}$$

Lemma 6.1. *For any $n \ge 1$ and any $0 \le t < \infty$*

$$\lim_{N\to\infty} E\left(\sup_{s\le t} |K^n_s(\phi) - K^{n,1}_s(\phi)| \right) = 0$$

The collision term $K^{n,1}_t(\phi)$ counts the number of occupied sites, $\nu_{n-1}(\beta)$, for each particle β who died before time t. Lemmas 5.1 and 5.3 imply that most of the the collision term comes from close relatives. To say how close these relatives are, we look at the conclusions of Lemma 5.3, and introduce a sequence of cutoffs τ_N defined by the requirements that:

(i) in $d > 2$, $N\tau_N \to \infty$ and $\tau_N \to 0$.

(ii) in $d = 2$, $\tau_N = 1/\log N$. (6.3)

Our next goal is to show that collisions involving two individuals more distantly related than τ_N in time can be ignored. To say this in symbols, we let

$$\mathrm{nbr}^m_{\beta,\gamma} = 1\{T_{\pi\gamma} < T_\beta \le T_\gamma,\ B^\gamma - B^\beta \in \mathscr{N}_N,\ \zeta^m_\gamma > T_{\pi\gamma}\}$$

We then can define the interference term for close relatives by

$$K^{n,2}_t(\phi) = \frac{1}{2N+\theta} \sum_\beta a^n_\beta(t)\phi(B^\beta)\, \frac{1}{\psi(N)} \sum_\gamma \mathrm{nbr}^{n-1}_{\beta,\gamma}\, 1(T_{\gamma\wedge\beta} > T_\beta - \tau_N)$$

where we recall that $\gamma \wedge \beta$ is the most recent common ancestor of γ and β and $T_{\gamma\wedge\beta} = -\infty$ if $\gamma_0 \ne \beta_0$.

Lemma 6.2. *For any $n \ge 1$ and $0 \le t < \infty$*

$$\lim_{N\to\infty} E\left(\sup_{s\le t} |K_s^{n,1}(\phi) - K_s^{n,2}(\phi)|\right) = 0$$

Most of the work for this has already been done in Section 5. However, notice that

$$\nu_{n-1}(\beta) \le \sum_\gamma \mathrm{nbr}_{\beta,\gamma}^{n-1}$$

since the left-hand side counts multiply occupied sites only once.

Our next step is to replace the requirements $\zeta_\beta^n > T_{\pi\beta}$ and $\zeta_\gamma^{n-1} > T_{\pi\gamma}$ by the condition

$$\zeta_\beta^n > T_\beta - \tau_N, \quad B^\beta \ne \Delta, \quad \zeta_\gamma^{n-1} > T_\beta - \tau_N, \quad B^\gamma \ne \Delta$$

and define

$$\begin{aligned} K_t^{n,3}(\phi) = \frac{1}{\psi(N)} \cdot \frac{1}{2N+\theta} \sum_\beta \phi(B^\beta) 1\{T_\beta \le t, \zeta_\beta^n > T_\beta - \tau_N\} \\ \times \sum_\gamma 1\{T_{\pi\gamma} < T_\beta \le T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N, \zeta_\gamma^{n-1} > T_\beta - \tau_N, \\ T_{\gamma\wedge\beta} > T_\beta - \tau_N\} \end{aligned} \tag{6.4}$$

Lemma 6.3. *For any $n \ge 1$ and $0 \le t < \infty$*

$$\lim_{N\to\infty} E\left(\sup_{s\le t} |K_s^{n,2}(\phi) - K_s^{n,3}(\phi)|\right) = 0$$

To prepare for the next step, we note that $X_t^1 \le X_t^2 \le X_t^0$. Thus, when $n = 1$ or 2

$$\begin{aligned} &\{\zeta_\beta^n > T_\beta - \tau_N,\ \zeta_\gamma^{n-1} > T_\beta - \tau_N,\ T_{\gamma\wedge\beta} > T_\beta - \tau_N\} \\ &\quad = \{\zeta_{\gamma\wedge\beta}^1 > T_\beta - \tau_N,\ T_{\gamma\wedge\beta} > T_\beta - \tau_N\} \\ &\quad = \{\zeta_\beta^1 > T_\beta - \tau_N,\ T_{\gamma\wedge\beta} > T_\beta - \tau_N\} \end{aligned}$$

At this point, we are finally ready to convert $K_t^{n,3}(\phi)$ into an integral. Let

$$F_\beta(r) = \frac{1}{\psi_0(N)} \sum_\gamma 1\{T_{\gamma\wedge\beta} > r - \tau_N, T_{\pi\gamma} < r \le T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\}$$

Here, $\psi_0(N) = \psi(N)/N$ is introduced to make this O(1). Note from the above that for $n = 1$ or 2

$$\begin{aligned} K_t^{n,3}(\phi) &= \frac{1}{N(2N+\theta)} \sum_\beta 1\{T_\beta \le t, \zeta_\beta^1 > T_\beta - \tau_N\} \phi(B^\beta) F_\beta(T_\beta) \\ &= K_t^{1,3}(\phi) \end{aligned}$$

This motivates the definition of

$$G_r^\tau(\phi) = \frac{1}{N} \sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, \zeta_\beta^1 > r - \tau_N\} \phi(B^\beta_{(r-\tau_N)^+})\, F_\beta(r) \quad (6.5)$$

Note that B^β has been replaced by the position of its family line at time $(r - \tau_n)^+$. More importantly $F_\beta(T_\beta)$ has turned into $F_\beta(r)$, and passing from the Poisson process to its compensator via Lemma 3.2 has removed the factor of $1/(2N + \theta)$.

Lemma 6.4. *For $n = 1$ and 2, and any $0 \le t < \infty$*

$$\lim_{N\to\infty} E\left(\sup_{s \le t} \left| K_s^{n,3}(\phi) - \int_0^s G_r^\tau(\phi)\, dr \right| \right) = 0$$

Let $\mathscr{A}(s) = \{\alpha : T_{\pi\alpha} < s \le T_\alpha, \zeta_\alpha^1 > s\}$ be the particles alive at time s in X^1. The contributions to the sum in $G_r^\tau(\phi)$ from the various particles in $\mathscr{A}(r - \tau_N)$, are independent, so it is natural to let

$$\{\alpha\}_r = \{\beta : \beta \ge \alpha, T_{\pi\beta} < r \le T_\beta, B^\beta \ne \Delta\} \quad (6.6)$$

be the set of descendants β of α that are alive at time r in the branching random walk, and use our new notation to write

$$G_r^\tau(\phi) = \frac{1}{N} \sum_{\alpha \in \mathscr{A}(r-\tau_N)} \phi(B^\alpha_{(r-\tau_N)^+}) Z_\alpha(r)$$

where $Z_\alpha(r) = \sum_{\beta \in \{\alpha\}_r} F_\beta(r)$. Comparing with

$$\begin{aligned} X_r^{1,\tau}(\phi) &\equiv \frac{1}{N} \sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, \zeta_\beta^1 > r - \tau_N\} \phi(B^\beta_{(r-\tau_N)^+}) \\ &= \frac{1}{N} \sum_{\alpha \in \mathscr{A}(r-\tau_N)} \phi(B^\alpha_{(r-\tau_N)^+}) |\{\alpha\}_r| \end{aligned}$$

suggests that we define $b_d^\tau = EZ_1(\tau_N)/E|\{1\}_{\tau_N}|$ (here $1 \in I$ labels the first individual in generation 0) and consider

$$G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi) = \frac{1}{N} \sum_{\alpha \in \mathscr{A}(r-\tau_N)} \phi(B^\alpha_{(r-\tau_N)^+}) \sum_{\beta \in \{\alpha\}_r} (F_\beta(r) - b_d^\tau) \quad (6.7)$$

By computing the variance of the last difference, we will conclude that

Lemma 6.5. *For any* $0 \le t < \infty$

$$\lim_{N\to\infty} E\left(\sup_{s \le t} \left| \int_0^s G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi)\, dr \right| \right) = 0$$

Recalling now that

$$X_r^1(\phi) = \frac{1}{N} \sum_\beta 1\{T_{\pi\beta} \le r < T_\beta, \zeta_\beta^1 > r\} \phi(B_r^\beta) \ ,$$

we see that it remains to remove the superscript τ's from $b_d^\tau X_r^{1,\tau}$ and complete the proof of the limit theorem for the collision term. This is a two-step procedure.

Lemma 6.6. $\lim_{N\to\infty} b_d^\tau = b_d$.

Lemma 6.7. *For any* $t < \infty$

$$\lim_{N\to\infty} E\left(\sup_{s \le t} \left| \int_0^s X_r^{1,\tau}(\phi) - X_r^1(\phi)\, dr \right| \right) = 0$$

Theorem 6.1, and hence Theorem 1, is an immediate consequence of Lemmas 6.1–6.7 (technically one also needs the trivial bound $\sup_N \sup_{r \le t} E(X_r^1(|\phi|)) < \infty$ from Lemma 2.9). The proofs of Lemmas 6.1–6.7 will keep us occupied until the end of Section 10. The rest of this section is devoted to proofs of the first two of these lemmas.

Proof of Lemma 6.1. Let $K_t^{n,0}$ be defined as K_t^n but with $\phi(B^\beta)$ in place of $\phi(B^\beta + W_\beta)$. Let

$$\eta_N = \sup\left\{ |\phi(x+y) - \phi(x)| : y \in [-N^{-1/2}, N^{-1/2}]^d \right\}$$

Any $\phi \in C_b^3$ is Lipschitz continuous, so $\eta_N \le CN^{-1/2} \to 0$ as $N \to \infty$. If we let

$$h_\beta^m = 1\{B^\beta + W_\beta \in \text{supp}(X^m_{T_\beta -})\}$$

be the indicator of the event that the support of X^m is hit by the birth at time T_β we can write

$$E\left(\sup_{t\le T}\left|K_t^n(\phi)-K_t^{n,0}(\phi)\right|\right)\le\frac{C}{N}E\left(\sum_\beta a_\beta^0(t)\eta_N h_\beta^0\right)$$
$$\le\eta_N\cdot C\big(X_0^0(1)+X_0^0(1)^2\big)\qquad(6.8)$$

by Lemma 2.5.

To estimate the difference between $K_t^{n,0}$ and $K_t^{n,1}$ now, we note that, as in the proof of Lemma 2.3,

$$E\left(h_\beta^{n-1}\,\middle|\,\mathscr{F}_{T_\beta-}\right)=\frac{\nu_{n-1}(\beta)}{\psi(N)},$$

so Lemma 3.5 implies that the difference

$$\begin{aligned}M_t&=\frac{K_t^{n,0}(\phi)}{1+\frac{\theta}{N}}-K_t^{n,1}(\phi)\\&=\frac{1}{2N+\theta}\sum_\beta a_\beta^n(t)\phi(B^\beta)\left\{h_\beta^{n-1}-\frac{\nu_{n-1}(\beta)}{\psi(N)}\right\}\end{aligned}\qquad(6.9)$$

is a martingale. To estimate the right hand side note that (i) the squares of the jumps of M_t are smaller than $N^{-2}\|\phi\|_\infty^2\{h_\beta^{n-1}-\nu_{n-1}(\beta)/\psi(N)\}^2$, (ii) $a_\beta^n(t)\le a_\beta^0(t)$, and (iii) since $h_\beta^{n-1}\in\{0,1\}$, $\mathrm{var}(h_\beta^{n-1})\le Eh_\beta^{n-1}$. So we have by the L^2 maximal inequality, (3.1),

$$\begin{aligned}E(\sup_{s\le t}M_s^2)&\le CE[M]_t\\&\le CN^{-2}E\left(\sum_\beta a_\beta^0(t)\,\|\phi\|_\infty^2\,h_\beta^{n-1}\right)\\&\le\frac{C}{N}\cdot\|\phi\|_\infty^2\left(X_0^0(1)+X_0^0(1)^2\right)\end{aligned}\qquad(6.10)$$

by Lemma 2.5. The desired conclusion now follows from (6.8)–(6.10) and the inequality

$$E\left(\sup_{s\le t}K_s^{n,0}\left|\frac{1}{1+\theta/N}-1\right|\right)\le\frac{C}{N}E(K_t^{n,0}(|\phi|))\to 0\text{ as }N\to\infty$$

(the last by Lemma 2.5 again) □

Proof of Lemma 6.2. Let

$$\nu_{m,\tau}(\beta)=|\{B^\gamma:T_{\pi\gamma}<T_\beta\le T_\gamma,\ B^\gamma-B^\beta\in\mathscr{N}_N,\ \zeta_\gamma^m>T_{\pi\gamma},$$
$$T_{\beta\wedge\gamma}>T_\beta-\tau_N\}|$$

be the number of neighbors of B^β occupied in X^m by close relatives of β at time T_β. To bridge the gap between $K_t^{n,1}$ and $K_t^{n,2}$ define

$$\hat{K}_t^{n,1}(\phi) = \frac{1}{2N+\theta} \sum_\beta a_\beta^n(t)\, \phi(B^\beta)\, \frac{\nu_{n-1,\tau}(\beta)}{\psi(N)}$$

Lemmas 5.1 and 5.3 imply that for any $t < \infty$

$$\lim_{N\to\infty} E\left(\sup_{s\le t} \left|K_s^{n,1}(\phi) - \hat{K}_s^{n,1}(\phi)\right|\right) = 0$$

To estimate the contribution to the collision term from births onto multiply occupied sites, we recall that $a_\beta^0(t) = \{T_\beta \le t,\, B^\beta \ne \Delta\}$ is the event that β was once alive in the branching random walk X^0 but died before time t, and let

$$J_1(t) = \frac{1}{N} \sum_\beta a_\beta^0(t)\, \frac{1}{\psi(N)} \sum_{\gamma:\gamma_0=\beta_0} \mathrm{nbr}^0_{\beta,\gamma} \times \sum_{\alpha\ne\gamma,\alpha_0=\gamma_0} 1\{T_{\pi\alpha} < T_\beta \le T_\alpha,\, B^\alpha = B^\gamma\} \qquad (6.11)$$

The motivation for this definition is that $|\hat{K}_t^{n,1}(\phi) - K_t^{n,2}(\phi)| \le J_1(t)\cdot\|\phi\|_\infty$. To check this observe that if there are $k \ge 2$ close relatives of B^β at one site B^γ neighboring B^β then the left-hand side contributes at most $\frac{(k-1)\phi(B^\beta)}{(2N+\theta)\psi(N)}$ but the right contributes at least $\frac{k(k-1)\|\phi\|_\infty}{N\psi(N)}$. The latter inequality comes from the fact that the definition of $J_1(t)$ in addition weakens the requirement of close relatives from that of having a recent common ancestor and of being alive in X^n or X^{n-1} to just being related and alive in X^0.

To estimate $EJ_1(t)$ we will have to sum and integrate over all the possibilities. The first step is to use the symmetry of (α, γ) and suppose without loss of generality that $\alpha \wedge \beta \le \gamma \wedge \beta$, i.e., that the α line did not split off from the β line after the γ line did. Just to keep on top of things the reader should note that there are two somewhat different sub-cases of this situation: (a) $\alpha \wedge \beta < \gamma \wedge \beta$, or (b) $\alpha \wedge \beta = \gamma \wedge \beta$. In words, (b) says that the most recent common ancestor of α and γ occurs after their common line of descent joins β.

To tackle (6.11) we begin with the inside sum and break things down according to the value of k so that $\gamma \wedge \alpha = \gamma|k = \alpha|k$, noting that the indicator functions involving β rule out $\alpha \wedge \gamma = \alpha$ or γ and so $k < |\alpha| \wedge |\gamma|$. Let $\mathscr{H}^\alpha_{\beta,\gamma}$ be the σ-field generated by all the branching events, and the random walk events for the lines β and γ only, but omitting the value of the jump $W_{\alpha\wedge\gamma}$ (which might have moved the α or the γ line). Then we have

$$P\big(B^\alpha = B^\gamma \,|\, \mathscr{H}^\alpha_{\beta,\gamma}\big) \le \frac{C}{\psi(N)}(|\alpha| - k)^{-d/2} \tag{6.12}$$

To prove this we use Lemma 4.2 to conclude that

$$\begin{aligned}&P(\sqrt{N}\{B^\alpha - B^\alpha(T_{\alpha|k})\} - \sqrt{N}\left\{B^\gamma - B^\gamma(T_{\gamma|k})\right\} \in [-1,1]^d |\mathscr{H}^\alpha_{\beta,\gamma})\\ &\quad\le C(|\alpha| - k)^{-d/2}\end{aligned}$$

and observe that the probability the missing $W_{\alpha\wedge\gamma}$ will have the exact value needed to make $B^\gamma = B^\alpha$ is at most $1/\psi(N)$.

In order to use (6.12), we want to take the conditional expectation of (6.11) with respect to $\mathscr{H}^\alpha_{\beta,\gamma}$. Unfortunately, $\text{nbr}^0_{\beta,\gamma}$ is not measurable with respect to $\mathscr{H}^\alpha_{\beta,\gamma}$. This problem is easy to fix. Let $\hat{B}^\beta$ (resp. $\hat{B}^\gamma$) be the position B^β (resp. B^γ) with the value of $W_{\alpha\wedge\gamma}$ subtracted if it appears in the sum. Clearly,

$$\text{nbr}^0_{\beta,\gamma} \le 1\{T_{\pi\gamma} < T_\beta \le T_\gamma,\ \sqrt{N}(\hat{B}^\beta - \hat{B}^\gamma) \in [-3,3]^d\} \tag{6.13}$$

and the right hand side is $\mathscr{H}^\alpha_{\beta,\gamma}-$measurable.

Modifying (6.11) using (6.13) then taking the expectation of the conditional expectation of the summands in (6.11) with respect to $\mathscr{H}^\alpha_{\beta,\gamma}$, we have

$$\begin{aligned}EJ_1(t) \le \frac{1}{N} E\sum_\beta a^0_\beta(t)\cdot\frac{1}{\psi(N)}&\\ \times \sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < T_\beta \le T_\gamma, \sqrt{N}(\hat{B}^\beta - \hat{B}^\gamma) \in [-3,3]^d\}&\\ \times \sum_{\alpha:\alpha\ne\gamma,\alpha_0=\gamma_0} \frac{C}{\psi(N)}(|\alpha| - |\alpha\wedge\gamma|)^{-d/2} 1(T_{\pi\alpha} < T_\beta \le T_\alpha)&\end{aligned} \tag{6.14}$$

To evaluate the inside sum we break things down according to the value of k so that $\gamma\wedge\alpha = \gamma|k = \alpha|k$, and the value of $\ell = |\alpha| - k - 1$. Since there are ℓ births in the α line after it splits from γ, there are 2^ℓ choices for α and each of them is alive with probability $\{(N+\theta)/(2N+\theta)\}^\ell$. Recalling there must also be exactly ℓ arrivals in the relevant rate $(2N+\theta)$ Poisson process, we arrive at

$$E\left(\sum_{\alpha:\alpha\ne\gamma,\alpha_0=\gamma_0} \frac{C}{\psi(N)}(|\alpha| - |\alpha\wedge\gamma|)^{-d/2} 1(T_{\pi\alpha} < T_\beta \le T_\alpha)\,\middle|\, \mathscr{H}^\alpha_{\beta,\gamma}\right)$$

$$\leq \sum_{k=0}^{|\gamma|-1} \sum_{\ell=0}^{\infty} 2^\ell \left(\frac{N+\theta}{2N+\theta}\right)^\ell \cdot \frac{C}{\psi(N)} (1+\ell)^{-d/2}$$
$$\times \exp[-(T_\beta - T_{\gamma|k})(2N+\theta)] \frac{[(T_\beta - T_{\gamma|k})(2N+\theta)]^\ell}{\ell!}$$
$$\leq \frac{C}{\psi(N)} \sum_{k=0}^{|\gamma|-1} \left(1 + (T_\beta - T_{\gamma|k})(2N+2\theta)\right)^{-d/2} \exp[\theta(T_\beta - T_{\gamma|k})]$$

by Lemma 4.3. Plugging this into (6.14) we may bound $EJ_1(t)$ by

$$\frac{C}{N\psi(N)^2} E \sum_\beta 1\{T_\beta \leq t\}$$
$$\times \sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < T_\beta \leq T_\gamma, \sqrt{N}(\hat{B}^\beta - \hat{B}^\gamma) \in [-3,3]^d\}$$
$$\cdot \sum_{k=0}^{|\gamma|-1} (1 + (T_\beta - T_{\gamma|k})(2N+2\theta))^{-d/2} \tag{6.15}$$

Using Lemma 3.2 to change from the Poisson jumps to their compensator (this introduces a factor of $2N+\theta$), and putting back in the variable $W_{\beta\wedge\gamma}$ left out of $\hat{B}^\beta$ and $\hat{B}^\gamma$, we see the above is at most

$$\frac{C}{\psi(N)^2} E \int_0^t \sum_\beta 1\{T_{\pi\beta} < r \leq T_\beta\}$$
$$\times \sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < r \leq T_\gamma, \sqrt{N}(B^\beta - B^\gamma) \in [-5,5]^d\}$$
$$\cdot \sum_{k=0}^{|\gamma|-1} (1 + (r - T_{\gamma|k})(2N+2\theta))^{-d/2}\, dr$$

If we sum over β first and condition on $\mathscr{H}_\gamma$, and then use (4.9), we see that the above is at most

$$\frac{C}{\psi(N)^2} E \int_0^t \sum_\gamma 1\{T_{\pi\gamma} < r \leq T_\gamma, B^\gamma \neq \Delta\}$$
$$\times \left\{\sum_{k=0}^{|\gamma|-1} (1 + (r - T_{\gamma|k})(2N+\theta))^{-d/2}\right\}^2 dr \tag{6.16}$$

Recalling there are $\{NX_0^0(1)\}$ choices for γ_0, then breaking things down according to the value of $m = |\gamma|$, and using the reasoning we applied to (6.14), we bound the above by

$$\frac{C}{\psi(N)^2}\{NX_0^0(1)\}\int_0^t \sum_{m=0}^{\infty} 2^m \left(\frac{N+\theta}{2N+\theta}\right)^m \times E\Bigg[1\{\Gamma_m < (2N+\theta)r < \Gamma_{m+1}\} \left(\sum_{k=0}^{m-1}(1+(2N+\theta)r-\Gamma_{k+1})^{-d/2}\right)^2\Bigg]dr$$

Using Lemma 4.6 and (4.7) it follows that the above is no more than

$$\frac{C}{\psi(N)^2}\{NX_0^0(1)\}\int_0^t \psi_0((2N+\theta)r)^2 dr \leq \frac{C}{\psi(N)^2}\{NX_0^0(1)\}\cdot t\psi_0((2N+\theta)t)^2$$

Recalling that $\psi_0((2N+\theta)t)/\psi(N) \leq C/N$ it follows that

$$EJ_1(t) \leq \frac{C}{N}X_0^0(1) \to 0$$

and the proof of Lemma 6.2 is complete. □

7. Proofs of Lemmas 6.4 and 6.5

For the moment we will skip Lemma 6.3, closing the loop with the proof of that result and the closely related Lemma 6.7 in Section 10.

Proof of Lemma 6.4. Recall that $\psi_0(N) = \psi(N)/N$,

$$F_\beta(r) = \frac{1}{\psi_0(N)}\sum_\gamma 1\{T_{\pi\gamma} < r \leq T_\gamma,\ T_{\gamma\wedge\beta} > r-\tau_N\} \cdot 1\{B^\gamma - B^\beta \in \mathcal{N}_N\}$$

$$G_r^\tau(\phi) = \frac{1}{N}\sum_\beta 1\{T_{\pi\beta} < r \leq T_\beta, \zeta_\beta^1 > r-\tau_N\}\phi(B^\beta_{(r-\tau_N)^+})F_\beta(r)$$

and the collision terms of interest, $K_t^{2,3} = K_t^{1,3}$. Our first observation is that

$$K_t^{1,3}(\phi) = \frac{1}{(2N+\theta)N}\sum_\beta 1\{T_\beta \leq t, \zeta_\beta^1 > T_\beta - \tau_N\}\phi(B^\beta)F_\beta(T_\beta)$$

is closely related to

$$G_r(\phi) = \frac{1}{N} \sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, \zeta^1_\beta > r - \tau_N\} \phi(B^\beta) F_\beta(r)$$

In particular, Lemma 3.2 (see also Lemma 3.4(a)) implies that

$$M_t = K^{1,3}_t(\phi) - \int_0^t G_r(\phi) dr \quad \text{is a martingale} \tag{7.1}$$

To bound the size of this martingale we note that the above Lemmas also imply

$$\langle M \rangle_t = \frac{1}{(2N+\theta)^2 N^2} \int_0^t \sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, \zeta^1_\beta > r - \tau_N\} \times \phi(B^\beta)^2 F_\beta(r)^2 \, (2N+\theta) dr$$

From the formula for $\langle M \rangle_t$, and very crude bounds, we get

$$E\langle M \rangle_T \le C\|\phi\|^2_\infty \int_0^T E\left(N^{-1} \sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, B^\beta \ne \Delta\} \times \left[N^{-1} \sum_\gamma 1\{\gamma_0 = \beta_0, T_{\pi\gamma} < r \le T_\gamma, B^\gamma \ne \Delta\} \right]^2 \right) dr$$

In the second equation, we have weakened the survival conditions to being alive in the branching random walk, so the above is at most

$$C\|\phi\|^2_\infty X^0_0(1) \int_0^T E\left(X^0_r(1)^3 \middle| \, X^0_0(1) = N^{-1}\delta_0 \right) dr \le C\|\phi\|^2_\infty X^0_0(1) \cdot T/N^{3/4} \tag{7.2}$$

The last inequality is immediate from Lemma 2.9 and Holder's inequality. The L^2 maximal inequality, (3.2), now shows that

$$E\left(\sup_{t \le T} M^2_t \right) \le C\|\phi\|^2_\infty X^0_0(1) T/N^{3/4} \ .$$

It remains to estimate the difference between the two G integrals. To this end we note that

$$E\big(|G^\tau_r(\phi) - G_r(\phi)|\big) \le \frac{1}{N} E\left(\sum_\beta 1\{T_{\pi\beta} < r \le T_\beta, \zeta^1_\beta > r - \tau_N\} \times \left|\phi(B^\beta_r) - \phi(B^\beta_{(r-\tau_N)^+})\right| \cdot F_\beta(r) \right) \tag{7.3}$$

Plugging in the definition of $F_\beta(r)$ then taking the conditional expectation with respect to $\mathscr{H}_\beta$ (recall its definition from the beginning of section 4), the above is no more than

$$\frac{C}{\psi(N)} E\left\{\sum_{\beta} 1\{T_{\pi\beta} < r \leq T_\beta\} \left|\phi(B_r^\beta) - \phi(B_{(r-\tau_N)^+}^\beta)\right| \right.$$
$$\left. \times E\left(\sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < r \leq T_\gamma, B^\gamma - B^\beta \in \mathcal{N}_N\} \middle| \mathcal{H}_\beta \right) \right\}$$

Note that we have eliminated $\zeta_\beta^1 > r - \tau_N$ and replaced $1\{T_{\gamma\wedge\beta} > r - \tau_N\}$ by $1(\gamma_0 = \beta_0)$. By (4.9), the above is bounded by

$$\frac{C}{\psi(N)} E\left\{\sum_{\beta} 1\{T_{\pi\beta} < r \leq T_\beta\} \left|\phi(B_r^\beta) - \phi(B_{(r-\tau_N)^+}^\beta)\right| \right.$$
$$\left. \times \sum_{k=0}^{|\beta|-1} \left(1 + (2N+\theta)(r - T_{\beta|k})\right)^{-d/2} \right\}$$

Our next step is to condition on

$$\mathcal{H}_\beta^{br} = \sigma(t_{\beta|m}, \delta_{\beta|m} : m < |\beta|) \vee \sigma(t_\beta) \;,$$

the information about the branching events in the family line of β plus the death time of β. Breaking things down according to the value of $m = |\beta|$ and then according to the values of the times $T_{\beta|k}$ the above is at most

$$\frac{C}{\psi(N)} \{N X_0^0(1)\} \sum_{m=1}^{\infty} 2^m \left(\frac{N+\theta}{2N+\theta}\right)^m E\left(\Sigma_m \Phi\right) \tag{7.4}$$

where $\Phi = E(|\phi(B_r^\beta) - \phi(B_{(r-\tau_N)^+}^\beta)||\mathcal{H}_\beta^{br})$,

$$\Sigma_m = 1\{\Gamma_m < (2N+\theta)r < \Gamma_{m+1}\} \sum_{k=0}^{m-1} (1 + \Gamma_m - \Gamma_{k+1})^{-d/2} \;,$$

and the Γ_m are the gamma random variables introduced in the proof of Lemma 4.5. Using the Cauchy-Schwarz inequality we conclude

$$E(\Sigma_m \Phi) \leq \left(E\Sigma_m^2\right)^{1/2} \left(E\Phi^2\right)^{1/2} \tag{7.5}$$

We have supposed that $\phi \in C_b^3$, and hence is Lipschitz continuous, so

$$E\Phi^2 \leq CE\left\{E\left(|B_r^\beta - B_{(r-\tau_N)^+}^\beta| \middle| \mathcal{H}_\beta^{br}\right)^2\right\} \leq CE\left|B_r^\beta - B_{(r-\tau_n)^+}^\beta\right|^2 \leq C\tau_N \tag{7.6}$$

Using (7.5) and (7.6) we see that (7.4) is smaller than

$$\frac{C\tau_N^{1/2}}{\psi(N)}\{NX_0^N(1)\}\sum_{m=1}^{\infty}2^m\left(\frac{N+\theta}{2N+\theta}\right)^m$$
$$\times\left\{E\left(1\{\Gamma_m<(2N+\theta)r<\Gamma_{m+1}\}\left(\sum_{k=0}^{m-1}(1+\Gamma_m-\Gamma_{k+1})^{-d/2}\right)^2\right)\right\}^{1/2}$$

Using Lemma 4.6 now, and (4.7) we see that the above is bounded by

$$\frac{C\tau_N^{1/2}}{\psi(N)}\{NX_0^0(1)\}\cdot I((2N+\theta)r)\le C\tau_N^{1/2}X_0^N(1)\to 0$$

This completes the proof of Lemma 6.4. □

Proof of Lemma 6.5. We begin by recalling formula (6.7):

$$G_r^\tau(\phi)-b_d^\tau X_r^{1,\tau}(\phi)=\frac{1}{N}\sum_{\alpha\in\mathscr{A}(r-\tau_N)}\phi(B^\alpha_{(r-\tau_N)^+})\sum_{\beta\in\{\alpha\}_r}(F_\beta(r)-b_d^\tau) \quad (7.7)$$

Plugging in the definition of $Z_\alpha(r)$ from Section 6 we get

$$G_r^\tau(\phi)-b_d^\tau X_r^{1,\tau}(\phi)=\frac{1}{N}\sum_{\alpha\in\mathscr{A}(r-\tau_N)}\phi(B^\alpha_{(r-\tau_N)^+})\left(Z_\alpha(r)-b_d^\tau|\{\alpha\}_r|\right) \quad (7.8)$$

If we condition on $\mathscr{F}_{r-\tau_N}$, then a simple argument using the Markov property and our basic independence assumptions shows that the individual summands in (7.8) are independent. The definition of b_d^τ now implies that their (conditional) means are 0. Here we use an obvious translation invariance to see that on $\{\alpha\in\mathscr{A}(r-\tau_N)\}$,

$$P((Z_\alpha(r),|\{\alpha\}_r|)\in\cdot|\mathscr{F}_{r-\tau_N})=P((Z_1(\tau_N),|\{1\}_{\tau_N}|)\in\cdot) \quad (7.9)$$

To show that the difference in (7.8) is small we will compute the variance of this random sum. For this it is clear from the above independence and equivalence in law that the following two lemmas will be needed.

Lemma 7.1. *There is a* $0<C<\infty$ *so that* $EZ_1^2(\tau_N)\le C(\tau N)$.

Lemma 7.2. *There is a* $0<C<\infty$ *so that for any* $s>0$, $E|\{1\}_s|^2\le C(1+sN)$.

The second result is a standard fact about critical branching processes and again is a corollary of Lemma 2.2 in Bramson, Durrett, and Swindle (1989). Before entering into the somewhat lengthy details of the proof of Lemma 7.1, let us check that it, and Lemma 6.6, will be enough to finish the proof of Lemma 6.5. Conditioning the sum in (7.8) on $\mathscr{F}_{r-\tau_N}$, we have from the above observations that if $r \geq \tau_N$ then

$$E\Big\{\big(G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi)\big)^2\Big\} \leq \frac{\|\phi\|_\infty^2}{N^2} E\big(|\mathscr{A}(r-\tau_N)| \cdot E\{(Z_1(\tau_N) - b_d^\tau|\{1\}_{\tau_N}|)^2\}\big) \quad (7.10)$$

Combining Lemmas 7.1, 7.2 and 6.6 (the latter to show that $\{b_d^\tau\}$ remains bounded as $N \to \infty$) and using the fact that we have chosen $N\tau_N \to \infty$ we have

$$E\{(Z_1(\tau_N) - b_d^\tau|\{1\}_{\tau_N}|)^2\} \leq CN\tau_N$$

so the expression is (7.10) is bounded by

$$\begin{aligned} C\tau_N\|\phi\|_\infty^2 E\left(N^{-1}|\mathscr{A}(r-\tau_N)|\right) &\leq C\tau_N\|\phi\|_\infty^2 E X_{r-\tau_N}^N(1) \\ &\leq C\tau_N\|\phi\|_\infty^2 X_0^N(1) \end{aligned}$$

by Lemma 2.9. From this it follows that

$$\begin{aligned} \int_{\tau_N}^T E|G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi)|dr &\leq \int_{\tau_N}^T \left(E|G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi)|^2\right)^{1/2} dr \\ &\leq C\tau_N^{1/2}\|\phi\|_\infty (X_0^N(1))^{1/2} \cdot T \qquad (7.11) \\ &\to 0 \text{ as } N \to \infty \end{aligned}$$

To handle the integral from 0 to τ_N we note that if $r \leq \tau_N$ then

$$\begin{aligned} G_r^\tau(\phi) &= N^{-1}\sum_\beta 1\{T_{\pi\beta} < r \leq T_\beta, \zeta_\beta^1 > r - \tau_N\}\phi(B_0^\beta)F_\beta(r) \\ X_r^{1,\tau}(\phi) &= N^{-1}\sum_\beta 1\{T_{\pi\beta} < r \leq T_\beta, \zeta_\beta^1 > r - \tau_N\}\phi(B_0^\beta) \end{aligned}$$

Using some trivial inequalities and then the definition of $F_\beta(r)$, we have

$$\begin{aligned}
&\int_0^{\tau_N} E|G_r^\tau(\phi) - b_d^\tau X_r^{1,\tau}(\phi)|\,dr \\
&\leq \|\phi\|_\infty \int_0^{\tau_N} (b_d^\tau E X_r^{1,\tau}(1) + E G_r^\tau(1))dr \\
&\leq \|\phi\|_\infty b_d^\tau \int_0^{\tau_N} E X_r^0(1) dr \\
&\quad + \int_0^{\tau_N} E\Bigg(\frac{\|\phi\|_\infty}{N\psi_0(N)} \sum_\beta 1\{T_{\pi\beta} < r \leq T_\beta\} \\
&\quad \cdot \sum_{\gamma:\gamma_0=\beta_0} 1\{T_{\pi\gamma} < r \leq T_\gamma, B^\gamma - B^\beta \in \mathscr{N}_N\}\Bigg) dr
\end{aligned}$$

Using Lemmas 2.9 and 6.6 on the first term and Lemma 4.4 and (4.7) on the second, the above is bounded by

$$C\|\phi\|_\infty X_0^N(1)\tau_N \to 0 \tag{7.12}$$

Combining (7.11) and (7.12) we see that the proof of Lemma 6.5 will be complete when we do the (independent!) proofs of Lemma 6.6 in Section 8 and Lemma 7.1 in Section 9. The latter result concerns the second moment $EZ_1(s)^2$, so we will first compute the mean $EZ_1(s)$, which is needed to prove Lemma 6.6.

8. Mean of the interference term

We claim that for $\alpha \in \mathscr{A}(r - \tau_N)$,

$$Z_\alpha(r) = \sum_{\beta \geq \alpha} \frac{1}{\psi_0(N)} \sum_{\gamma \geq \alpha} 1\{T_{\pi\beta} < r \leq T_\beta, T_{\pi\gamma} < r \leq T_\gamma, B^\beta - B^\gamma \in \mathscr{N}_N\} \tag{8.1}$$

To see this note that for $\alpha \in \mathscr{A}(r - \tau_N)$ and $\beta \geq \alpha$ the condition $T_{\gamma\wedge\beta} > r - \tau_N$ (appearing in the definition of $F_\beta(r)$) holds iff $\gamma \geq \alpha$ (which appears in the definition of $Z_\alpha(r)$). For this equivalence, observe that $T_{\gamma\wedge\beta} \geq r - \tau_N$ implies $T_{\gamma\wedge\beta} > T_{\pi\alpha}$ and since $\gamma \wedge \beta$ and α are both ancestors of β this forces $\gamma \wedge \beta \geq \alpha$ and so $\gamma \geq \alpha$. The converse implication is obvious and (8.1) now follows from the definitions of $Z_\alpha(r)$ and $F_\beta(r)$. If $\beta = \gamma$, then $B^\beta - B^\gamma \notin \mathscr{N}_N$ and so (8.1) with $r = \tau_N$ (clearly $1 \in \mathscr{A}(0)$) implies that

$$\begin{aligned}
&\psi_0(N) E Z_1(\tau_N) \\
&= E\left(\sum_{\beta \geq 1} \sum_{\gamma \geq 1, \gamma \neq \beta} 1\{T_{\pi\beta} < \tau_N \leq T_\beta, T_{\pi\gamma} < \tau_N \leq T_\gamma, B^\beta - B^\gamma \in \mathscr{N}_N\} \right)
\end{aligned}$$

Note that for a fixed β and γ in the above sum if $\alpha = \beta \wedge \gamma$ then the following are mutually independent σ-fields:

$$\mathscr{H}_\alpha,\ \sigma(t_{\alpha'} : \alpha < \alpha' \leq \beta),\ \sigma(t_{\alpha'} : \alpha < \alpha' \leq \gamma),$$
$$\sigma(\bar{B}^\gamma - \bar{B}^\beta),\ \sigma(\delta_{\alpha'} : \alpha \leq \alpha' \leq \beta \text{ or } \alpha < \alpha' \leq \gamma)$$

Breaking things down according to the value of $\alpha = \beta \wedge \gamma$, using the above independence, and conditioning on $\mathscr{B} = \sigma(t_{\alpha'}, \delta_{\alpha'} : \alpha' \in I) \wedge \mathscr{H}_\alpha$ and then on $\mathscr{H}_\alpha$ shows that

$$\begin{aligned}\psi_0(N) E Z_1(\tau_N) = E\Bigg\{&\sum_{k=0}^\infty \sum_{\alpha:|\alpha|=k,\,\alpha\geq 1} 1\{T_\alpha < \tau_N, B^\alpha \neq \Delta\}\\ &\cdot 2 \cdot \sum_{\ell=0}^\infty \sum_{m=0}^\infty \sum_{\substack{\beta>\alpha,|\beta|=k+1+\ell\\ \beta_{k+1}=0}} \sum_{\substack{\gamma>\alpha,|\gamma|=k+1+m\\ \gamma_{k+1}=1}} \left(\frac{N+\theta}{2N+\theta}\right)^{1+\ell+m}\\ &\cdot P(T_{\pi\beta} - T_\alpha < \tau_N - T_\alpha \leq T_\beta - T_\alpha | \mathscr{H}_\alpha)\\ &\cdot P(T_{\pi\gamma} - T_\alpha < \tau_N - T_\alpha \leq T_\gamma - T_\alpha | \mathscr{H}_\alpha)\\ &\cdot P\big(N^{1/2}(\bar{B}^\gamma - \bar{B}^\beta) \in [-1,1]^d - \{0\}\big)\Bigg\}\end{aligned} \tag{8.2}$$

Starting at the bottom of (8.2), if W^N is uniform on $N^{1/2} \cdot \mathscr{N}_N$, V_n^N is an independent random walk that with probability 1/2 stays put and with probability 1/2 takes a step uniform on $N^{1/2} \cdot \mathscr{N}_N$, then

$$\begin{aligned}&P\big(N^{1/2}(\bar{B}^\gamma - \bar{B}^\beta) \in [-1,1]^d - \{0\}\big)\\ &\quad = P\big(W^N + V_{\ell+m}^N \in [-1,1]^d - \{0\}\big)\end{aligned}$$

Combine this with the usual Poisson process formulas for the probability β and γ are alive at time τ_N, to equate (8.2) to

$$\begin{aligned}E\Bigg\{&\sum_{k=0}^\infty \sum_{\alpha:|\alpha|=k,\,\alpha\geq 1} 1\{T_\alpha < \tau_N, B^\alpha \neq \Delta\} \cdot 2 \cdot \sum_{\ell=0}^\infty \sum_{m=0}^\infty 2^{\ell+m} \left(\frac{N+\theta}{2N+\theta}\right)^{1+\ell+m}\\ &\times \frac{((2N+\theta)(\tau_N - T_\alpha))^{\ell+m}}{\ell!\,m!} e^{-2(2N+\theta)(\tau_N - T_\alpha)}\\ &\times P\big(W^N + V_{\ell+m}^N \in [-1,1]^d - \{0\}\big)\Bigg\}\end{aligned}$$

Recall $\epsilon_N = \theta/(2N+\theta)$. Changing variables from (ℓ, m) to (n, m) where $n = \ell + m$, gives

$$(1+\epsilon_N)\, E\left\{\sum_{k=0}^{\infty} \sum_{\alpha:|\alpha|=k,\,\alpha\geq 1} 1\{T_\alpha < \tau_N,\, B^\alpha \neq \Delta\} \cdot \sum_{n=0}^{\infty}\sum_{m=0}^{n} \left(\frac{2N+2\theta}{2N+\theta}\right)^n \right.$$
$$\times \frac{n!}{(n-m)!\,m!} \cdot \frac{((2N+\theta)(\tau_N - T_\alpha))^n}{n!} e^{-2(2N+\theta)(\tau_N - T_\alpha)}$$
$$\left. \times P(W^N + V_n^N \in [-1,1]^d - \{0\}) \right\} \tag{8.3}$$

Summing $\binom{n}{m}$ over m from 0 to n gives 2^n. A little arithmetic turns the sum over n into

$$e^{2\theta(\tau_N - T_\alpha)} \sum_{n=0}^{\infty} \frac{(4(N+\theta)(\tau_N - T_\alpha))^n}{n!}$$
$$\times\, e^{-4(N+\theta)(\tau_N - T_\alpha)} P(W^N + V_n^N \in [-1,1]^d - \{0\}) \tag{8.4}$$

Let $\pi(u)$ be a Poisson random variable with mean u that is independent of W_N and $\{V_n^N : n \geq 0\}$, and let

$$h_N(u) = P\big(W^N + V_{\pi(u)}^N \in [-1,1]^d - \{0\}\big)$$

Using our new notation we can write (8.4) as

$$e^{2\theta(\tau_N - T_\alpha)} h_N(4(N+\theta)(\tau_N - T_\alpha)) \ .$$

Plugging this into (8.3) we see that (8.3) equals

$$(1+\epsilon_N)\, E\ \left\{\sum_{\alpha\geq 1} 1\{T_\alpha < \tau_N,\, B^\alpha \neq \Delta\} \cdot e^{2\theta(\tau_N - T_\alpha)}\right.$$
$$\left. h_N(4(N+\theta)(\tau_N - T_\alpha))\right\} \tag{8.5}$$

Using Lemma 3.2 and Lemma 3.4(a) for integrability, we may convert the above to

$$(1+\epsilon_N)\, E\int_0^{\tau_N} \left\{\sum_{\alpha\geq 1} 1\{T_{\pi\alpha} < r \leq T_\alpha,\, B^\alpha \neq \Delta\}\right.$$
$$\left. \cdot e^{2\theta(\tau_N - r)} h_N(4(N+\theta)(\tau_N - r)) \cdot (2N+\theta) \right\} dr$$

Summing over α gives all the individuals alive in the branching process at time r, so using Lemma 2.9, we have shown

$$\psi_0(N)EZ_1(\tau_N)$$
$$= (1+\epsilon_N)\, e^{\theta\tau_N} \int_0^{\tau_N} e^{2\theta(\tau_N - r)} h_N(4(N+\theta)(\tau_N - r)) \cdot (2N+\theta) dr$$

We can simplify our calculations by noting that ($\sim$ indicates the ratio approaches 1 as $N \to \infty$)

$$\psi_0(N)EZ_1(\tau_N) \sim \int_0^{\tau_N} h_N\big(4(N+\theta)(\tau_N - r)\big) \cdot (2N+2\theta) dr$$
$$= \frac{1}{2} \int_0^{4(N+\theta)\tau_N} h_N(s) ds \tag{8.6}$$

where in the second step we have changed variables $s = 4(N+\theta)(\tau_N - r)$.

Let W be uniform over $[-1,1]^d$ and $V_{\pi(s)}$ be an independent continuous time random walk that at rate 1/2 takes a step uniform on $[-1,1]^d$. Elementary weak convergence arguments show that

Lemma 8.1. *If $s_N \to s < \infty$ then as $N \to \infty$,*

$$h_N(s_N) \to h(s) = P\big(W + V_{\pi(s)} \in [-1,1]^d\big)$$

We now wish to interchange the limit as $N \to \infty$ and the integral over s in (8.6). In $d > 2$ this is easy to justify. Lemma 2.4 in Bramson, Durrett, and Swindle (1989) gives

$$P\big(x + V^N_{\pi(t)} \in [-1,1]^d\big) \le C(1+t)^{-d/2} \tag{8.7}$$

Since the right hand side is independent of x, the same bound holds when W^N is put in place of x on the left. This gives us the domination we need to let $N \to \infty$ in (8.6) and conclude that if $N\tau_N \to \infty$ and $\tau_N \to 0$ then

$$\lim_{N\to\infty} \psi_0(N)EZ_1(\tau_N) = \frac{1}{2} E \int_0^{\infty} P\big(W + V_{\pi(s)} \in [-1,1]^d\big)\, ds$$

In $d > 2$, $\psi(N) \sim 2^d N$ so $\psi_0(N) \to 2^d$. For the right-hand side, we note that the continuous time random walk $V_{\pi(s)}$ stays in each state for an exponential amount of time with mean 1/2 before moving, so, recalling the definition of $\{U_n\}$ in Section 1 prior to Theorem 1, we can rewrite the last formula as

$$\lim_{N\to\infty} EZ_1(\tau_N) = 2^{-d} \sum_{n=1}^{\infty} P(U_n \in [-1,1]^d) = b_d \tag{8.8}$$

Things are a little more delicate in $d = 2$ since the limiting integral is divergent. Fortunately, much of the work has been done in Lemma 4.6 of Bramson, Durrett, and Swindle (1989). Here $|B|$ denotes the Lebesgue measure of B.

Lemma 8.2. *If $s \to \infty$, $x_s/(s/2)^{1/2} \to x$, and $N \to \infty$ then for any Borel set B with $|B| < \infty$ and $|\partial B| = 0$*

$$(s/2)^{d/2} P(V^N_{\pi(s)} \in x_s + B) \to |B|n(x)$$

where $n(x) = (2\pi/3)^{-d/2} \exp(-3|x|^2/2)$ is the normal density with variance $1/3$.

The $t/2$ comes from the fact that in Bramson, Durrett, and Swindle (1989), what they call X^M_t takes jumps at rate 1 while our $V^N_{\pi(s)}$ takes jumps at rate $1/2$, so we need to set $n = s/2$, $M = N$ in their Lemma 4.6.

Lemma 8.3. *If $s_N \to \infty$ then $h_N(s_N)/h(s_N) \to 1$.*

Proof. Use the classical local central limit theorem to see that

$$(s_N/2)^{d/2} h(s_N) \to 2^d n(0) \tag{8.9}$$

(Problem 1 in Section 10.4 of Breiman (1968) and a simple calculation will suffice.) Note that by conditioning on $W^N \in [1, 1]^d$ and using Lemma 8.2 and (8.7) to integrate out the conditioning, we get

$$(s_N/2)^{d/2} h_N(s_N) \to 2^d n(0) \enspace .$$

The result follows. □

Combining this with Lemma 8.1, one can conclude easily that

Lemma 8.4. *Let $\eta > 0$. If N is large then $h_N(s)/h(s) \in [1 - \eta, 1 + \eta]$ for all $s \geq 0$.*

Proof. Let $\eta > 0$ and suppose that there is a sequence of exceptions s_{N_k} to the inequality. There is either a subsequence converging to a finite limit or the sequence converges to ∞. In the first case we contradict Lemma 8.1, in the second we contradict Lemma 8.3. □

From Lemma 8.4 it is immediate that

$$\int_0^{4(N+\theta)\tau_N} h_N(s)ds \sim \int_0^{4(N+\theta)\tau_N} h(s)ds \enspace , \tag{8.10}$$

that is, the ratio approaches 1 as $N \to \infty$. To compute the right-hand side, we use (8.9) (with $d = 2$) to get

$$h(s) \sim (s/2)^{-1} \cdot 4 \cdot (2\pi/3)^{-1} = \frac{12}{\pi s} \text{ as } s \to \infty$$

In $d = 2$, $\tau_N = 1/\log N$ and $\psi_0(N) \sim 4 \log N$, so (8.6), (8.10), the above asymptotic estimate, and the trivial bound $h(s) \leq 1$ imply that as $N \to \infty$

$$EZ_1(\tau_N) \sim \frac{1}{4 \log N} \cdot \frac{1}{2} \int_{\log \log N}^{4(N+\theta)/\log N} \frac{12}{\pi s} \, ds \to \frac{3}{2\pi} \tag{8.11}$$

Formulas (8.11) and (8.8) give the asymptotic behavior of $EZ_1(\tau_N)$ for $d = 2$ and $d > 2$, respectively. To complete the proof of Lemma 6.6 now, we note that by Lemma 2.9 and the fact that $\tau_N \to 0$,

$$E|\{1\}_{\tau_N}| = e^{\theta \tau_N} \to 1 \tag{8.12}$$

as $N \to \infty$. Therefore from (8.8), (8.11) and (8.12) we see that as $N \to \infty$,

$$b_d^\tau = \frac{EZ_1(\tau_N)}{E|\{1\}_{\tau_N}|} \to b_d$$

and Lemma 6.6 is proved.

9. Second moment of the interference term

In this section we will prove Lemma 7.1. We will use $\beta \approx s$ to indicate that $T_{\pi\beta} < s \leq T_\beta$, $B^\beta \neq \Delta$. Using this in (8.1) we can write

$$\psi_0(N)^2 Z_1(\tau_N)^2 = \sum_{\beta_1,\beta_2,\beta_3,\beta_4} \left(\prod_{i=1}^4 1\{\beta_i \approx \tau_N\} \right) 1\{B^{\beta_1} - B^{\beta_2} \in \mathcal{N}_N, B^{\beta_3} - B^{\beta_4} \in \mathcal{N}_N\} \tag{9.1}$$

where each β_i has $\beta_i(0) = 1$. To suppress nuisance terms later it is useful to note:

(i) Since all the β_i are alive at time s we cannot have $\beta_i < \beta_j$.

(ii) Since $0 \notin \mathcal{N}_N$ we must have $\beta_1 \neq \beta_2$ and $\beta_3 \neq \beta_4$.

(iii) From (i), (ii), and $\beta_i(0) = 1$ it follows that $|\beta_i| \geq 1$ for all i.

There are several cases in the estimation of (9.1) depending on the relative relationship of the the β_i. To sort these out we need some notation. For a finite set Λ of possible individuals let

$$\tau(\Lambda;\beta)=\max\{|\gamma\wedge\beta|:\gamma\in\Lambda,\gamma_0=\beta_0\}$$

with $\max\emptyset=-\infty$. In words, $\tau(\Lambda;\beta)$ is the number of the last generation in which β had an ancestor in common with some individual in Λ. For $j=1,2,3$, let R_j be the contribution to the sum in (9.1) from $\beta_1,\beta_2,\beta_3,\beta_4$ with $\tau(\{\beta_1,\beta_2,\beta_3\},\beta_4)=|\beta_j\wedge\beta_4|$. The contributions we have defined overlap but the terms in the sum are nonnegative so

$$\psi_0(N)^2E\left(Z_1(\tau_N)^2\right)\le\sum_{i=1}^{3}ER_i=2ER_2+ER_3 \tag{9.2}$$

where in the second step we have used symmetry to conclude $ER_1=ER_2$.

To estimate the right-hand side of (9.2) we have to do each of the four sums in (9.1). To structure the proof we will divide this section into the corresponding subsections.

a. Sum over β_4. For $j=1,2,3$, we let

$$R_j(\beta_1,\beta_2,\beta_3)$$
$$=\sum_{\beta_4}1\left\{\tau(\{\beta_1,\beta_2,\beta_3\};\beta_4)=|\beta_j\wedge\beta_4|\right\}\cdot 1(B^{\beta_3}-B^{\beta_4}\in\mathcal{N}_N)$$

and note that conditioning on $\mathcal{H}_{123}=\mathcal{H}_{\beta_1}\vee\mathcal{H}_{\beta_2}\vee\mathcal{H}_{\beta_3}$,

$$ER_j=\sum_{\beta_1,\beta_2,\beta_3}\left(\prod_{i=1}^{3}1\{\beta_i\approx\tau_N\}\right)1\{B^{\beta_1}-B^{\beta_2}\in\mathcal{N}_N\}$$
$$\times E(R_j(\beta_1,\beta_2,\beta_3)|\mathcal{H}_{123}) \tag{9.3}$$

Breaking things down according to the value of $k=|\beta_j\wedge\beta_4|$ and using Lemma 4.2, we have that $E(R_j(\beta_1,\beta_2,\beta_3)|\mathcal{H}_{123})-1$ is at most

$$C\sum_{k=0}^{|\beta_j|-1}\sum_{\beta_4,|\beta_4|>k}1\left\{\tau(\{\beta_1,\beta_2,\beta_3\};\beta_4)=|\beta_j\wedge\beta_4|=k\right\}\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta_4|-k-1}$$
$$\times(|\beta_4|-k)^{-d/2}\cdot P\left(T_{\pi\beta_4}<\tau_N\le T_{\beta_4}\middle|\mathcal{H}_{123}\right) \tag{9.4}$$

Here the 1 corresponds to the term $\beta_4=\beta_j$ (which contributes if $j\ne 3$) and we have used (i) to justify $|\beta_4|>k$. If $u(\gamma)=(2N+\theta)(\tau_N-T_\gamma)$, then on $\{\beta_j\approx\tau_N\}$ the above is no more than

$$C\sum_{k=0}^{|\beta_j|-1}\sum_{\beta_4,|\beta_4|>k}1\left\{\tau(\{\beta_1,\beta_2,\beta_3\};\beta_4)=|\beta_j\wedge\beta_4|=k\right\}\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta_4|-k-1}$$

$$\times(|\beta_4| - k)^{-d/2} \cdot e^{-u(\beta_j|k)} \frac{u(\beta_j|k)^{|\beta_4|-k-1}}{(|\beta_4| - k - 1)!}$$

Letting $\ell = |\beta_4| - k - 1$ and taking into account the number of possible β_4, we may bound the above on $\{\beta_j \approx \tau_N\}$ by

$$\begin{aligned} C \sum_{k=0}^{|\beta_j|-1} \sum_{\ell=0}^{\infty} \left(\frac{2N+2\theta}{2N+\theta}\right)^{\ell} (1+\ell)^{-d/2} \cdot e^{-u(\beta_j|k)} \frac{u(\beta_j|k)^{\ell}}{\ell!} \\ \leq C \sum_{k=0}^{|\beta_j|-1} e^{\theta(\tau_N - T_{\beta_j|k})}(1 + u(\beta_j|k))^{-d/2} \\ \leq Ce^{\theta\tau_N} H(\beta_j, \tau_N) \leq CH(\beta_j, \tau_N) \end{aligned} \tag{9.5}$$

by Lemma 4.3 and a definition given after (4.9). We can combine (9.4) and (9.5) to get

$$1(\beta_j \approx \tau_N)E\left(R_j(\beta_1, \beta_2, \beta_3)\middle|\mathscr{H}_{123}\right) \leq 1 + CH(\beta_j, \tau_N) \tag{9.6}$$

Note. It is tempting to use $H(\beta_j, \tau_N) \leq H(\beta_j)$ and $H(\beta_j) \geq 1$ to simplify the right hand side to $CH(\beta_j)$ but that upper bound does not work well in (9.8).

b. Sum on β_3. Using (9.2), (9.3), and (9.6), then conditioning on $\mathscr{H}_{12} = \mathscr{H}_{\beta_1} \vee \mathscr{H}_{\beta_2}$

$$\begin{aligned} &\psi_0(N)^2 E Z_1(\tau_N)^2 \\ &\leq CE\Bigg\{\sum_{\beta_1}\sum_{\beta_2} 1\left(\beta_1 \approx \tau_N, \beta_2 \approx \tau_N, B^{\beta_1} - B^{\beta_2} \in \mathscr{N}_N\right) \\ &\qquad \times \Bigg[E\bigg(\sum_{\beta_3} 1(\beta_3 \approx \tau_N)[1 + CH(\beta_3, \tau_N)]\bigg|\mathscr{H}_{12}\bigg) \\ &\qquad + [1 + CH(\beta_2, \tau_N)] \cdot E\bigg(\sum_{\beta_3} 1(\beta_3 \approx \tau_N)\bigg|\mathscr{H}_{12}\bigg)\Bigg]\Bigg\} \end{aligned} \tag{9.7}$$

Using symmetry we can replace $1(\beta_3 \approx \tau_N)$ in the first sum by

$$1\{\beta_3 \approx \tau_N, \tau(\{\beta_1, \beta_2\}; \beta_3) = |\beta_2 \wedge \beta_3|\}$$

and put another factor of 2 into the C. To deal with the right-hand side of (9.7) we let

$$R_j(\beta_1,\beta_2) = \sum_{\beta_3} 1\Big(\beta_3 \approx \tau_N,\, \tau(\{\beta_1,\beta_2\};\beta_3) = |\beta_j \wedge \beta_3|\Big) \cdot H(\beta_3)$$

$$Q_j(\beta_1,\beta_2) = \sum_{\beta_3} 1\Big(\beta_3 \approx \tau_N,\, \tau(\{\beta_1,\beta_2\};\beta_3) = |\beta_j \wedge \beta_3|\Big)$$

for $j = 1, 2$. Separating out the possibility of $\beta_3 = \beta_2$ first, and plugging in the definition of $H(\beta_3, \tau_N)$, we see that $E(R_2(\beta_1,\beta_2)|\mathscr{H}_{12})$ is bounded by

$$\begin{aligned} H(\beta_2) + \sum_{k=0}^{|\beta_2|-1} \sum_{\beta_3} 1(\tau(\{\beta_1,\beta_2\};\beta_3) = |\beta_2 \wedge \beta_3| = k) \left(\frac{N+\theta}{2N+\theta}\right)^{|\beta_3|-k-1} \\ \times \Bigg\{ E\Bigg[1\{T_{\pi\beta_3} < \tau_N \le T_{\beta_3}\} \sum_{j=k}^{|\beta_3|-1} [1 + u(\beta_3|j)]^{-d/2} \Bigg| \mathscr{H}_{12} \Bigg] \\ + E\Bigg[1\{T_{\pi\beta_3} < \tau_N \le T_{\beta_3}\} \Bigg| \mathscr{H}_{12} \Bigg] \cdot \sum_{j=0}^{k-1} [1 + u(\beta_2|j)]^{-d/2} \Bigg\} \end{aligned} \tag{9.8}$$

where we have used the fact that for $j < k$, $\beta_3|j = \beta_2|j$.

Introducing $\ell = |\beta_3| - k - 1$, changing variables $i = j - k$, and using our standard gamma random variables Γ_m defined in the proof of Lemma 4.5, the first term in the set braces in (9.8) is at most

$$E\left(1\{\Gamma_\ell < (2N+\theta)(\tau_N - T_{\beta_3|k}) < \Gamma_{\ell+1}\} \sum_{i=0}^{\ell} (1 + \Gamma_\ell - \Gamma_i)^{-d/2} \right) \tag{9.9}$$

Using the usual Poisson reasoning with the trivial bound $k \le |\beta_2|$, and recalling the definition of $H(\beta_2)$, we see that the second term in the set braces in (9.8) is bounded by

$$e^{-u(\beta_2|k)} \frac{u(\beta_2|k)^\ell}{\ell!} H(\beta_2) \tag{9.10}$$

Using (9.9) and (9.10) in (9.8), and counting the number of β_3's for each ℓ, we have

$$\begin{aligned} E(R_2(\beta_1,\beta_2)|\mathscr{H}_{12}) \ \le H(\beta_2) + \sum_{k=0}^{|\beta_2|-1} \sum_{\ell=0}^{\infty} \left(\frac{2N+2\theta}{2N+\theta}\right)^\ell \\ \times \Bigg\{ E\Bigg(1\{\Gamma_\ell < (2N+\theta)(\tau_N - T_{\beta_3|k}) < \Gamma_{\ell+1}\} \\ \times \sum_{i=0}^{\ell} (1 + \Gamma_\ell - \Gamma_i)^{-d/2} \Bigg) + e^{-u(\beta_2|k)} \frac{u(\beta_2|k)^\ell}{\ell!} H(\beta_2) \Bigg\} \end{aligned} \tag{9.11}$$

Recalling the notation $e_\ell(u)$ from (4.12) we see that the first term in set braces in (9.11) is at most $2e_\ell((2N+\theta)(\tau_N - T_{\beta_3|k}))$. Here the factor 2 is used to handle the $i = 0$ term which doesn't appear in the sum defining $e_\ell(u)$. Doing the sum over ℓ now and using the above and Lemma 4.5, with the trivial bound $T_{\beta_3|k} \geq 0$, we see that the first term in set braces in (9.11) when summed over ℓ and k contributes at most

$$C|\beta_2| \cdot I((2N+\theta)\tau_N) \tag{9.12}$$

The second term in set braces in (9.11) when summed contributes at most

$$\begin{aligned}&\sum_{k=0}^{|\beta_2|-1}\sum_{\ell=0}^{\infty}(1+\epsilon_N)^\ell e^{-u(\beta_2|k)}\frac{u(\beta_2|k)^\ell}{\ell!}H(\beta_2)\\ &\quad= \sum_{k=0}^{|\beta_2|-1}\exp(\epsilon_N u(\beta_2|k))\, H(\beta_2) \leq C|\beta_2|H(\beta_2)\end{aligned} \tag{9.13}$$

since $\epsilon_N = \theta/(2N+\theta)$ and $u(\beta_2|k) \leq (2N+\theta)\tau_N$. Using (9.12) and (9.13) in (9.11), then recalling $|\beta_i| \geq 1$ by (iii), we have

$$E(R_2(\beta_1,\beta_2)|\mathscr{H}_{12}) \leq C|\beta_2| \cdot \left\{I((2N+\theta)\tau_N) + H(\beta_2)\right\} \tag{9.14}$$

Our next step is to consider Q_j. With $H(\beta_3, \tau_N)$ in R_j being replaced by 1 in Q_j, the analysis is much easier. Imitating (9.8) we write for $j = 1, 2$

$$\begin{aligned}E(Q_j(\beta_1,\beta_2)|\mathscr{H}_{12}) \leq 1 + &\sum_{k=0}^{|\beta_j|-1}\sum_{\beta_3} 1\Big\{\tau(\{\beta_1,\beta_2\};\beta_3) = |\beta_j \wedge \beta_3| = k\Big\}\\ &\times\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta_3|-k-1} P(T_{\pi\beta_3} < \tau_N \leq T_{\beta_3}|\mathscr{H}_{12})\end{aligned}$$

Introducing $\ell = |\beta_3| - k - 1$, we see the above is at most

$$1 + \sum_{k=0}^{|\beta_j|-1}\sum_{\ell=0}^{\infty}\left(\frac{2N+2\theta}{2N+\theta}\right)^\ell e^{-u(\beta_j|k)}\frac{u(\beta_j|k)^\ell}{\ell!} \leq C|\beta_j| \tag{9.15}$$

since $\theta u(\beta_j|k)/(2N+\theta) \leq \theta\tau_N$, and (iii) tells us that $|\beta_j| \geq 1$.

(9.14) and (9.15) handle the first term in square brackets in (9.7). To take care of the second term there, we note that

$$E\left(\sum_{\beta_3} 1(\beta_3 \approx \tau_N)\middle|\mathscr{H}_{12}\right) \leq E\left(Q_1(\beta_1,\beta_2) + Q_2(\beta_1,\beta_2)\middle|\mathscr{H}_{12}\right) \;,$$

so using (9.15), $H(\beta_2, \tau_N) \le H(\beta_2)$ and $H(\beta_2) \ge 1$ we have

$$E\left(\sum_{\beta_3} 1(\beta_3 \approx \tau_N)[1 + CH(\beta_2, \tau_N)] \middle| \mathscr{H}_{12}\right) \le C(|\beta_1| + |\beta_2|) \cdot H(\beta_2) \tag{9.16}$$

Using (9.14) and (9.16), we see that (9.7) is bounded by

$$CE\left\{\sum_{\beta_1}\sum_{\beta_2} 1\left(\beta_1 \approx \tau_N, \beta_2 \approx \tau_N, B^{\beta_1} - B^{\beta_2} \in \mathscr{N}_N\right)\right.$$
$$\left. \times \left[|\beta_2| \cdot \left\{I((2N+\theta)\tau_N) + H(\beta_2)\right\} + (|\beta_1| + |\beta_2|) \cdot H(\beta_2)\right]\right\} \tag{9.17}$$

Having summed over β_4 and then β_3, our third step is to
c. Sum over β_1. If we condition on $\mathscr{H}_2 = \mathscr{H}_{\beta_2}$ in (9.17) then we will be left with two types of terms:

$$R(\beta_2) = \sum_{\beta_1} 1(\beta_1 \approx \tau_N, B^{\beta_1} - B^{\beta_2} \in \mathscr{N}_N) \cdot |\beta_1|$$
$$Q(\beta_2) = \sum_{\beta_1} 1(\beta_1 \approx \tau_N, B^{\beta_1} - B^{\beta_2} \in \mathscr{N}_N)$$

Using our new notation, $E((9.17)|\mathscr{H}_2)$ is no more than

$$CE\left\{\sum_{\beta_2} 1(\beta_2 \approx \tau_N)\big[H(\beta_2) \cdot E(R(\beta_2)|\mathscr{H}_2)\right.$$
$$\left. + |\beta_2| \cdot \big(I((2N+\theta)\tau_N) + H(\beta_2)\big) \cdot E(Q(\beta_2)|\mathscr{H}_2)\big]\right\} \tag{9.18}$$

(4.9) implies that

$$E(Q(\beta_2)|\mathscr{H}_2) \le CH(\beta_2) \tag{9.19}$$

To cope with the extra factor of $|\beta_1|$ in $R(\beta_2)$, we note that, adapting the proof of (4.8), one can easily show

$$E(R(\beta_2)|\mathscr{H}_2) \le C \sum_{k=0}^{|\beta_2|-1} \sum_{\ell=0}^{\infty} (\ell + 1 + k) \cdot (\ell + 1)^{-d/2}$$
$$\times \left(\frac{2N + 2\theta}{2N + \theta}\right)^{\ell} e^{-u(\beta_2|k)} \frac{u(\beta_2|k)^{\ell}}{\ell!}$$

Dividing the sum into (a) $\ell \geq k$, where $\ell+1+k \leq 2(\ell+1)$, and (b) $\ell < k$, where $\ell+1+k \leq 2k \leq 2|\beta_2|$, and then using Lemma 4.3 on each piece, we see that the above is less than or equal to

$$C \sum_{k=0}^{|\beta_2|-1} (1+u(\beta_2|k))^{1-d/2} + C|\beta_2| \sum_{k=0}^{|\beta_2|-1} (1+u(\beta_2|k))^{-d/2}$$

The second sum is at most $H(\beta_2)$. For the first we use $(1+u(\beta_2|k))^{1-d/2} \leq 1$, which holds in $d \geq 2$, and $H(\beta_2) \geq 1$ to get

$$E(R(\beta_2)|\mathscr{H}_2) \leq C|\beta_2|H(\beta_2) \tag{9.20}$$

At last, we are ready for the fourth and final step.
d. Sum on β_2. Using (9.19) and (9.20) we see that the mean value of (9.18) is at most

$$CE\left\{ \sum_{\beta_2} 1(\beta_2 \approx \tau_N)|\beta_2| \left(H(\beta_2) I((2N+\theta)\tau_N) + H(\beta_2)^2 \right) \right\} \tag{9.21}$$

Breaking things down according to the value of $\ell = |\beta_2|$ we may bound the above by (recall the notation e_ℓ and g_ℓ from (4.12) and (4.19))

$$\begin{aligned} C \sum_{\ell=1}^{\infty} 2^\ell \left(\frac{N+\theta}{2N+\theta} \right)^\ell \ell \cdot E \Bigg\{ & 1(\Gamma_\ell < (2N+\theta)\tau_N < \Gamma_{\ell+1}) \\ & \times \Bigg[I((2N+\theta)\tau_N) \left(\sum_{k=0}^{\ell-1} (1+\Gamma_\ell - \Gamma_{k+1})^{-d/2} \right) \\ & + \left(\sum_{k=0}^{\ell-1} (1+\Gamma_\ell - \Gamma_{k+1})^{-d/2} \right)^2 \Bigg] \Bigg\} \\ = C \sum_{\ell=1}^{\infty} (1+\epsilon_N)^\ell \ell \big[& I((2N+\theta)\tau_N) e_\ell((2N+\theta)\tau_N) \\ & + g_\ell((2N+\theta)\tau_N) \big] \end{aligned}$$

Let $v = (2N+\theta)\tau_N$ and write $\sum_{\ell \in A}$ for the above sum when ℓ is restricted to A. If $\ell \leq 3v$, then $\ell(1+\epsilon_N)^\ell \leq Cv$, and so Lemmas 4.5 and 4.6 imply

$$\sum_{\ell \leq 3v} \leq C \cdot v \cdot I(v)^2 \leq CN\tau_N \cdot I(N)^2 \tag{9.22}$$

Using the trivial bound $\Gamma_\ell - \Gamma_k \geq 0$, we see that

$$\sum_{\ell > 3v} \leq C \sum_{\ell > 3v} (1+\epsilon_N)^\ell \cdot \ell \cdot e^{-v} \frac{v^\ell}{\ell!} [I(v)\ell + \ell^2] \leq CI(v) \sum_{\ell > 3v} a_\ell(v)$$

where $a_\ell(v) = (1+\epsilon_N)^\ell e^{-v} v^\ell \ell^3/\ell!$. If N is large, $a_{\ell+1}(v)/a_\ell(v) \leq 1/2$ for all $\ell > 3v$, so

$$\sum_{\ell > 3v} a_\ell(v) \leq 2a_{3v}(v) \to 0$$

exponentially fast as $v \to \infty$ by standard large deviations estimates for the Poisson distribution. (See e.g., page 82 of Durrett (1995a).) From the last result it follows that

$$I(v) \sum_{\ell > 3v} a_\ell(v) \leq C \tag{9.24}$$

Combining (9.24) with (9.22) and recalling $v = (2N+\theta)\tau_N$, it follows that (9.21), and hence $\psi_0(N)^2 E(Z_1(\tau_N)^2)$ (recall (9.7)), is at most

$$C(1 + I(N)^2 \tau_N N) \leq CI(N)^2 \tau_N N$$

Now use (4.7) to obtain the conclusion of Lemma 7.1. □

10. Proofs of Lemmas 6.3 and 6.7

First consider Lemma 6.3. In the definition of $a^n_\beta(t)$ we implicitly used the fact that $\zeta^n_\beta = T_{\beta|k}$ for some k to see that

$$1(\zeta^n_\beta > T_{\pi\beta}) = 1(\zeta^n_\beta \geq T_\beta)$$

The same reasoning for γ shows that

$$1(T_{\pi\gamma} < T_\beta \leq T_\gamma, \zeta^{n-1}_\gamma > T_{\pi\gamma}) = 1(T_{\pi\gamma} < T_\beta \leq T_\gamma, \zeta^{n-1}_\gamma \geq T_\beta)$$

Hence in the definition of $K^{n,2}$ we can replace the conditions $\zeta^n_\beta > T_{\pi\beta}$, $\zeta^{n-1}_\gamma > T_{\pi\gamma}$ with $\zeta^n_\beta \geq T_\beta,\ \zeta^{n-1}_\gamma \geq T_\beta$. Taking differences and replacing γ by δ we therefore have

$$\begin{aligned} &\sup_{s \leq t} |K^{n,2}_s(\phi) - K^{n,3}_s(\phi)| \\ &\quad \leq \frac{1}{(2N+\theta)\psi(N)} \sum_{\beta,\delta} |\phi(B^\beta)| \end{aligned}$$

$$\times 1\left\{T_\beta \le t, T_{\pi\delta} < T_\beta \le T_\delta, B^\delta - B^\beta \in \mathcal{N}_N, T_{\delta\wedge\beta} > T_\beta - \tau_N\right\}$$
$$\times\left[1\{\zeta_\beta^n > T_\beta - \tau_N, \zeta_\delta^{n-1} > T_\beta - \tau_N\} - 1\left\{\zeta_\beta^n \ge T_\beta, \zeta_\delta^{n-1} \ge T_\beta)\right\}\right]$$

Use Lemma 3.2 (and Lemma 3.4 (a) for integrability) to bound the mean value of the above by

$$\frac{\|\phi\|}{\psi(N)} E \int_0^t \sum_{\beta,\delta} 1\{T_{\pi\beta} < r \le T_\beta, T_{\pi\delta} < r \le T_\delta, B^\delta - B^\beta \in \mathcal{N}_N,$$
$$T_{\delta\wedge\beta} > r - \tau_N\}\left[1\{\zeta_\beta^n > r - \tau_N, \zeta_\delta^{n-1} > r - \tau_N\}\right.$$
$$\left.-1\{\zeta_\beta^n \ge r, \zeta_\delta^{n-1} \ge r\}\right]dr$$

which is a nicer form since β and δ play exactly symmetric roles. Subtracting and adding $1(\zeta_\beta^n \ge r, \zeta_\delta^{n-1} > r - \tau_N)$ and using symmetry, we see that in order to demonstrate Lemma 6.3, it is enough to establish for all $m \ge 0$, that

$$\frac{1}{\psi(N)} E \int_0^t \sum_{\beta,\delta} 1 \;\left\{T_{\pi\beta} < r \le T_\beta, T_{\pi\delta} < r \le T_\delta, B^\delta - B^\beta \in \mathcal{N}_N,\right.$$
$$\left. T_{\delta\wedge\beta} > r - \tau_N, \zeta_\beta^m \in [r - \tau_N, r]\right\} dr \to 0 \quad (10.1)$$

as $N \to \infty$. Since $T_\beta \ge r$ and $B^\beta \ne \Delta$ imply $\zeta_\beta^0 > r$ a.s. this is trivial for $m = 0$.

To start to work on $m \ge 1$ we introduce

$$I_\beta(r) = \frac{1}{\psi_0(N)} \sum_{\delta:\delta_0=\beta_0} 1\{T_{\pi\delta} < r \le T_\delta, B^\delta - B^\beta \in \mathcal{N}_N\}$$

Here we divide by $\psi_0(N)$ to make $I_\beta(r)$ be $O(1)$. Note that $T_{\delta\wedge\beta} > r - \tau_N$ implies $\delta_0 = \beta_0$ and so to establish (10.1) it is enough to show

Lemma 10.1. *For any $m \ge 1$, as $N \to \infty$*

$$\frac{1}{N} E \int_0^t \sum_\beta 1 \;\{T_{\pi\beta} < r \le T_\beta, \zeta_\beta^m \in [r - \tau_N, r], B^\beta \ne \Delta\}$$
$$\times \left(I_\beta(r) + 1\right) dr \to 0$$

Here, the $+1$ is not needed for Lemma 6.3, but is included for the

Proof of Lemma 6.7. Recalling the definitions given in Section 6, and using the fact that our test functions $\phi \in C_b^3$ are Lipschitz continuous, and $T_{\pi\beta} < r$ iff $T_{\pi\beta} \le r$ a.s., we have

$$E\left(\int_0^t |X_r^{1,\tau}(\phi) - X_r^1(\phi)|\,dr\right)$$
$$\le \|\phi\| \int_0^t E\Big(\frac{1}{N}\sum_\beta 1\{T_{\pi\beta} < r < T_\beta,\, \zeta_\beta^1 \in [r-\tau_N, r],\, B^\beta \ne \Delta\}\Big)\,dr$$
$$+C\int_0^t E\Big(\frac{1}{N}\sum_\beta 1\{T_{\pi\beta} < r < T_\beta,\, B^\beta \ne \Delta\}|B_r^\beta - B^\beta_{(r-\tau_N)^+}| \wedge 1\Big)dr$$

The first term tends to 0 by Lemma 10.1. For the second term, condition on $\mathscr{F}_{(r-\tau_N)^+}$ and use the Markov property and Lemma 2.9 with $\phi(x) = |x| \wedge 1$ to see that if $B^N(t)$ is the continuous time random walk in Lemma 2.9 then the second term is at most

$$CX_0^N(1)\int_0^t e^{\theta r}\,dr\; E|B^N(\tau_N) - B^N(0)| \le CX_0^N(1)\sqrt{\tau_N}$$

Combine these last two observations with the fact that $\tau_N \to 0$ to complete the proof. □

It remains then to do the

Proof of Lemma 10.1. Now if β is alive in the branching process at time r but has $\zeta_\beta^m \in [r-\tau_N, r]$), then there is a $i < |\beta|$ so that

$$T_{\beta|i} \ge r - \tau_N, \quad B^{\beta|i+1} \in \text{supp}(X^{m-1}(T_{\beta|i}-)), \quad e_{\beta|i} = \beta(i+1)$$

In words the last condition says that at time $T_{\beta|i}$ the β line experienced the dispersal event and collided with a particle already present in X^{m-1}. Let γ denote the index of one of the particles with which β collides at time $T_{\beta|i}$. Using $D_{\beta,i}$ as short hand for the awkward $e_{\beta|i} = \beta(i+1)$, and reading the symbol as "there was a displacement in the family line of β at the death of $\beta|i$," we can bound the quantity of interest in Lemma 10.1 by

$$\frac{1}{N}\int_0^t E\sum_\beta 1\{T_{\pi\beta} < r \le T_\beta,\, B^\beta \ne \Delta\}\sum_{i=0}^{|\beta|-1} 1\{D_{\beta,i},\, T_{\beta|i} \ge r - \tau_N\}$$
$$\times \sum_\gamma 1\{T_{\pi\gamma} < T_{\beta|i} < T_\gamma,\, B^{\beta|i+1} = B^\gamma\}\big[I_\beta(r) + 1\big]\,dr \qquad (10.2)$$

The first step in bounding this is to let $\mathscr{H}_{\beta,\gamma} = \mathscr{H}_\beta \vee \mathscr{H}_\gamma$, recall

$$H(\alpha) = \sum_{j=0}^{|\alpha|-1}\big[1 + (2N+\theta)(T_{\pi\alpha} - T_{\alpha|j})\big]^{-d/2} \quad ,$$

and generalize the proof of (4.9) to show

Lemma 10.2. *On* $\{T_{\pi\beta} < r\} \cap \{T_{\pi\gamma} < r\}$,

$$E\big(I_\beta(r)\big|\,\mathcal{H}_{\beta,\gamma}\big) \le \frac{C}{\psi_0(N)}\,(H(\beta)+H(\gamma))$$

Proof. Recall that $\alpha \approx r$ means $T_{\pi\alpha} < r \le T_\alpha$ and $B^\alpha \ne \Delta$. On $\{T_{\pi\beta} < r,\ T_{\pi\gamma} < r\}$ for $\alpha = \beta$ or γ, set

$$I_{\beta,\alpha}(r) = \frac{1}{\psi_0(N)} \sum_{\delta:\delta(0)=\beta(0)} 1\ \Big\{\delta \approx r,\ B^\delta - B^\beta \in \mathcal{N}_N,\ \tau(\{\beta,\gamma\};\delta) = |\alpha\wedge\delta|\Big\}$$

Since either β or γ must split off from δ last (it can be a tie if the δ lineage branches off before the β and γ lines separate) we have

$$E\left(I_\beta(r)\big|\,\mathcal{H}_{\beta,\gamma}\right) \le E\left(I_{\beta,\beta}(r)\ |\mathcal{H}_{\beta,\gamma}\right) + E\left(I_{\beta,\gamma}(r)\big|\,\mathcal{H}_{\beta,\gamma}\right)$$

where the second term can only contribute if $\beta_0 = \gamma_0$. Thus it suffices to show that for $\alpha = \beta$ and $\alpha = \gamma$ that

$$E\left(I_{\beta,\alpha}(r)\ |\mathcal{H}_{\beta,\gamma}\right) \le \frac{C}{\psi_0(N)}\cdot H(\alpha)$$

Break things down according to the value of $\tau(\{\beta,\gamma\},\delta) = |\alpha\wedge\delta| = k \in \{0,\ldots,|\alpha|\}$, isolate the case $\delta = \alpha$ first, and then observe that when $\delta \ne \alpha$, $\delta \approx r$ and $T_{\pi\alpha} < r$ imply $|\delta| > k$. This gives

$$\begin{aligned}
&\psi_0(N)\cdot E\left(I_{\beta,\alpha}(r)\big|\,\mathcal{H}_{\beta,\gamma}\right)\\
&\le 1 + \sum_{k=0}^{|\alpha|}\sum_{\delta,|\delta|>k} 1\,(\tau(\{\beta,\gamma\};\delta) = k = |\alpha\wedge\delta|)\\
&\quad\times E\big(1(T_{\pi\delta} < r \le T_\delta,\ B^\delta \ne \Delta)\\
&\qquad P\big(B^\beta - B^\delta \in \mathcal{N}_N\big|\,\mathcal{H}_{\beta,\gamma} \vee \mathcal{H}^{br}_\delta\big)\big|\,\mathcal{H}_{\beta,\gamma}\big)
\end{aligned}$$

where $\mathcal{H}^{br}_\delta = \sigma(t_{\delta|m}, \delta_{\delta|m} : m < |\delta|) \vee \sigma(t_\delta)$ is the σ-field generated by the branching events in the family line of δ. Let $u(\alpha|k,r) = (2N+\theta)(r - T_{\alpha|k})$ and note that for the $k = |\alpha|$ term to contribute in the above sum we must have $T_\alpha < r$. Using Lemma 4.2 now and setting $\ell = |\delta| - k - 1$, we may bound the above by

$$1 + \sum_{k=0}^{|\alpha|}\sum_{\ell=0}^{\infty}(1+\epsilon_N)^\ell \cdot C(1+\ell)^{-d/2}\cdot e^{-u(\alpha|k,r)^+}\frac{u(\alpha|k,r)^{+\ell}}{\ell!}$$

Using Lemma 4.3 now with the trivial inequalities $u(\alpha|k,r)^+ \geq (2N+\theta)(T_{\pi\alpha} - T_{\alpha|k})^+$ (from our original hypothesis on $T_{\pi\alpha}$), and $u(\alpha|k,r)^+ \leq (2N+\theta)r$, the above is bounded by

$$1 + Ce^{\epsilon_N(2N+\theta)r}\left[\sum_{k=0}^{|\alpha|-1}\Big(1+(2N+\theta)(T_{\pi\alpha}-T_{\alpha|k})\Big)^{-d/2}+1\right]$$

and the desired result follows from the definition of $H(\alpha)$. □

Conditioning (10.2) on $\mathscr{H}_{\beta,\gamma}$, using Lemma 10.2, and throwing away the event $D_{\beta|i}$, we see that (10.2) is at most

$$\begin{aligned}&\frac{C}{N}\int_0^t E\sum_\beta 1(T_{\pi\beta} < r \leq T_\beta,\, B^\beta \neq \Delta)\sum_{i=0}^{|\beta|-1} 1(T_{\beta|i} > T_{\pi\beta} - \tau_N)\\ &\quad\times\sum_\gamma 1(T_{\pi\gamma} < T_{\beta|i} < T_\gamma,\, B^{\beta|i+1} = B^\gamma)\\ &\quad\times\big[1+\psi_0(N)^{-1}(H(\beta)+H(\gamma))\big]\,dr\end{aligned}$$

Doing the integral over r now and recalling $t_\beta = T_\beta - T_{\pi\beta}$ we may bound the above by

$$\begin{aligned}&\frac{C}{N}E\sum_\beta t_\beta 1(T_{\pi\beta} \leq t,\, B^\beta \neq \Delta)\sum_{i=0}^{|\beta|-1} 1(T_{\beta|i} > T_{\pi\beta} - \tau_N)\\ &\quad\times\sum_\gamma 1(T_{\pi\gamma} < T_{\beta|i} < T_\gamma,\, B^{\beta|i+1} = B^\gamma)\\ &\quad\times\big[1+\psi_0(N)^{-1}(H(\beta)+H(\gamma))\big]\end{aligned}$$

To get rid of the t_β, we condition on the random variables generating $\mathscr{H}_{\beta,\gamma}$, but without t_β, and note that if γ is not a descendant of β (which we may assume in light of the condition $T_{\pi\gamma} < T_{\beta|i} < T_\gamma$) then this information is independent of t_β. We thus may conclude that the above is no more than

$$\begin{aligned}&\frac{C}{N^2}E\Bigg\{\sum_\beta 1\left(T_{\pi\beta} \leq t,\, B^\beta \neq \Delta\right)\sum_{i=0}^{|\beta|-1} 1\left(T_{\beta|i} > T_{\pi\beta} - \tau_N\right)\\ &\quad\times\sum_\gamma 1\left(T_{\pi\gamma} < T_{\beta|i} < T_\gamma,\, B^{\beta|i+1} = B^\gamma\right)\\ &\quad\quad\times[1+\psi_0(N)^{-1}(H(\beta)+H(\gamma))]\Bigg\}\end{aligned} \tag{10.3}$$

Let $\mathscr{H}^i_{\beta,\gamma}$ be $\mathscr{H}_{\beta,\gamma}$ but without the information about the value of $W^{\beta|i}$. Since at most one of the $\psi(N)$ values of W can make a perfect hit of B^γ, and none will hit unless the source of the birth, $B^{\beta|i}$, is a neighbor of B^γ, we have

$$P\big(B^{\beta|i+1}=B^\gamma \,\big|\, \mathscr{H}^i_{\beta,\gamma}\big)=\frac{1}{\psi(N)}\cdot 1\left(B^{\beta|i}-B^\gamma\in\mathscr{N}_N\right)$$

To use this, we condition (10.3) on $\mathscr{H}^i_{\beta,\gamma}$ to see that it is bounded by

$$\frac{C}{N^2}E\Bigg\{\sum_\beta 1(T_{\pi\beta}\le t,\,B^\beta\ne\Delta)\sum_{i=0}^{|\beta|-1}1(T_{\beta|i}>T_{\pi\beta}-\tau_N)$$
$$\times\sum_\gamma 1(T_{\pi\gamma}<T_{\beta|i}<T_\gamma)\cdot\frac{1}{\psi(N)}1(B^{\beta|i}-B^\gamma\in\mathscr{N}_N)$$
$$\times[1+\psi_0(N)^{-1}(H(\beta)+H(\gamma))]\Bigg\}$$

Our next step is to break things down according to the value of $k=|\beta\wedge\gamma|$ (set $|\emptyset|=-1$) which must be less than i, and condition on $\mathscr{T}_\beta\vee\mathscr{T}_\gamma$, where $\mathscr{T}_\alpha=\sigma(t_{\alpha|m}:m\le|\alpha|)$ is the information about the branching times in the line of α. Using Lemma 4.2, the last display is bounded by

$$\frac{C}{N^2\psi(N)}E\Bigg\{\sum_\beta 1(T_{\pi\beta}\le t)\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta|}\sum_{i=0}^{|\beta|-1}1(T_{\beta|i}>T_{\pi\beta}-\tau_N)$$
$$\times\sum_{k=-1}^{i-1}\;\sum_{\gamma,|\gamma\wedge\beta|=k<|\gamma|}1(T_{\pi\gamma}<T_{\beta|i}<T_\gamma)\cdot(|\gamma|-k)^{-d/2}$$
$$\times\left(\frac{N+\theta}{2N+\theta}\right)^{|\gamma|-k-1}\cdot\left[1+\frac{H(\beta)+H(\gamma)}{\psi_0(N)}\right]\Bigg\}\tag{10.4}$$

To bound the right-hand side of (10.4), we will handle the $H(\beta)/\psi_0(N)$ and $1+H(\gamma)/\psi_0(N)$ terms in the last square brackets separately, and call the resulting sums (10.4a) and (10.4b). For the first we will condition on $\mathscr{H}_\beta$ and break things down according to the value of $\ell=|\gamma|-k-1$. We set $T_{\beta|k}=0$ if $k=-1$. Then

$$E\left(\sum_{\gamma:|\gamma\wedge\beta|=k<|\gamma|}1\big(T_{\pi\gamma}<T_{\beta|i}<T_\gamma\big)(|\gamma|-k)^{-d/2}\left(\frac{N+\theta}{2N+\theta}\right)^{|\gamma|-k-1}\Bigg|\mathscr{H}_\beta\right)$$
$$\le\sum_{\ell=0}^{\infty}(1+\epsilon_N)^\ell(\ell+1)^{-d/2}P\big(\Gamma_\ell<(2N+\theta)(T_{\beta|i}-T_{\beta|k})<\Gamma_{\ell+1}\big)$$

By the usual Poisson reasoning, this equals

$$\sum_{\ell=0}^{\infty}(1+\epsilon_N)^{\ell}(\ell+1)^{-d/2}\cdot e^{-(2N+\theta)(T_{\beta|i}-T_{\beta|k})}\frac{[(2N+\theta)(T_{\beta|i}-T_{\beta|k})]^{\ell}}{\ell!}$$
$$\leq Ce^{\theta T_{\beta|i}}\left(1+(2N+\theta)(T_{\beta|i}-T_{\beta|k})\right)^{-d/2}$$

by Lemma 4.3. Plugging this bound in we see that (10.4a) is no more than

$$\frac{C}{N^2\psi(N)}E\Bigg\{\sum_{\beta}1(T_{\pi\beta}\leq t)\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta|}\sum_{i=0}^{|\beta|-1}1(T_{\beta|i}>T_{\pi\beta}-\tau_N)$$
$$\times\frac{H(\beta)}{\psi_0(N)}\cdot\sum_{k=-1}^{i-1}\left(1+(2N+\theta)(T_{\beta|i}-T_{\beta|k})\right)^{-d/2}\Bigg\}$$

Breaking things down according to the value of $\ell=|\beta|-1$, writing out the definition of $H(\beta)$, and introducing our standard gamma random variables, we bound the above by

$$\frac{C}{N^2\psi(N)}\cdot\{NX_0^N(1)\}E\sum_{\ell=0}^{\infty}1\left(\Gamma_{\ell+1}\leq(2N+\theta)t\right)(1+\epsilon_N)^{\ell+1}$$
$$\times\sum_{i=0}^{\ell}1(\Gamma_{i+1}>\Gamma_{\ell+1}-(2N+\theta)\tau_N)\sum_{k=-1}^{i-1}(1+\Gamma_{i+1}-\Gamma_{k+1})^{-d/2}$$
$$\times\frac{1}{\psi_0(N)}\left\{\sum_{j=0}^{\ell}\left(1+\Gamma_{\ell+1}-\Gamma_{j+1}\right)^{-d/2}\right\}\tag{10.5}$$

Bounding (10.5) is a Poisson process exercise, which we will attend to later, so we turn now to the other piece of (10.4). To handle (10.4b), we begin by interchanging the order of summations to get

$$\frac{C}{N^2\psi(N)}E\Bigg\{\sum_{\gamma}1(T_{\pi\gamma}\leq t)\left[1+\frac{H(\gamma)}{\psi_0(N)}\right]$$
$$\times\sum_{k=-1}^{|\gamma|-1}(|\gamma|-k)^{-d/2}\left(\frac{N+\theta}{2N+\theta}\right)^{|\gamma|}\sum_{\beta,|\beta\wedge\gamma|=k<|\beta|}1(T_{\pi\beta}\leq t)$$
$$\times\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta|-k-1}\sum_{i=k+1}^{|\beta|-1}1(T_{\pi\gamma}<T_{\beta|i}<T_{\gamma},T_{\beta|i}>T_{\pi\beta}-\tau_N)\Bigg\}\tag{10.6}$$

Conditioning on $\mathscr{H}_{\gamma}$, and introducing $\ell=|\beta|-k-1$, $j=i-k$, $u_k=(2N+\theta)(T_{\pi\gamma}-T_{\beta|k})$, and $v_k=(2N+\theta)(T_{\gamma}-T_{\beta|k})$ (if $k=-1$ then $T_{\beta|k}=0$ as above), we have

$$E\Bigg(\sum_{\beta,|\beta\wedge\gamma|=k<|\beta|} 1(T_{\pi\beta\le t})\left(\frac{N+\theta}{2N+\theta}\right)^{|\beta|-k-1} \times \sum_{i=k+1}^{|\beta|-1} 1(T_{\pi\gamma} < T_{\beta|i} < T_{\gamma}, T_{\beta|i} > T_{\pi\beta}-\tau_N)\Big|\,\mathscr{H}_{\gamma}\Bigg)$$
$$\le \sum_{\ell=1}^{\infty}(1+\epsilon_N)^{\ell}\sum_{j=1}^{\ell} P(\Gamma_{\ell}\le(2N+\theta)t, u_k<\Gamma_j<v_k, \Gamma_{\ell}-\Gamma_j\le(2N+\theta)\tau_N)$$

Introducing $x=\Gamma_j$ and $y=\Gamma_{\ell}-\Gamma_j$, and putting the case $j=\ell$ into the second term, we may bound the previous display by

$$\iint 1(x+y\le(2N+\theta)t,\ u_k<x<v_k,\ y\le(2N+\theta)\tau_N) \times\left[\sum_{\ell=2}^{\infty}(1+\epsilon_N)^{\ell}\cdot\sum_{j=1}^{\ell-1}\frac{x^{j-1}}{(j-1)!}\frac{y^{\ell-j-1}}{(\ell-j-1)!}e^{-x-y}\right]dx\,dy + \int 1(x\le(2N+\theta)t, u_k<x<v_k)\sum_{\ell=1}^{\infty}(1+\epsilon_N)^{\ell}\frac{x^{\ell-1}}{(\ell-1)!}e^{-x}\,dx \tag{10.7}$$

Using the identity

$$\sum_{j=1}^{\ell-1}\frac{x^{j-1}}{(j-1)!}\frac{y^{\ell-j-1}}{(\ell-j-1)!}=\frac{(x+y)^{\ell-2}}{(\ell-2)!}$$

we can rewrite the double sum in square brackets in (10.7) as

$$\sum_{\ell=2}^{\infty}(1+\epsilon_N)^{\ell}\frac{(x+y)^{\ell-2}}{(\ell-2)!}e^{-x-y}\le Ce^{\epsilon_N(x+y)}\le C$$

if $(x+y)\le(2N+\theta)t$. Evaluating the single sum in the same way, and throwing away the first restriction on x we see that (10.7) is at most

$$C\iint 1(u_k<x<v_k, y\le(2N+\theta)\tau_N)dxdy + C\int 1(u_k<x<v_k)dx$$

Recall that $v_k-u_k=(2N+\theta)t_{\gamma}$, where $t_{\gamma}=T_{\gamma}-T_{\pi\gamma}$ is the lifetime of γ, and that our choices in (6.3) imply that $(2N+\theta)\tau_N\to\infty$, to bound the above by

$$C(2N+\theta)\tau_N\cdot(2N+\theta)t_{\gamma}$$

Using the last inequality, we see that (10.6) is bounded by

$$\frac{C}{N^2\psi(N)} E\Bigg\{ \sum_{\gamma} 1(T_{\pi\gamma} \le t) \left[1 + \frac{H(\gamma)}{\psi_0(N)}\right] \times \sum_{k=-1}^{|\gamma|-1} (|\gamma| - k)^{-d/2} \left(\frac{N+\theta}{2N+\theta}\right)^{|\gamma|} (2N+\theta)\tau_N \cdot (2N+\theta)t_\gamma \Bigg\} \quad (10.8)$$

As before one can condition on everything but t_γ and use $E(2N+\theta)t_\gamma = 1$ to get rid of that term. Breaking things down according to the value of $\ell = |\gamma| - 1$, separating out the contribution from $|\gamma| = 0$ and noting that $H(\gamma) = 0$ if $|\gamma| = 0$, using $\sum_{k=-1}^{\ell}(1 + \ell - k)^{-d/2} \le C\psi_0(\ell)$, and filling in the definition of $H(\gamma)$, we see the above is bounded by

$$\frac{C\tau_N}{N\psi(N)} \cdot \{NX_0^0(1)\} \cdot \Bigg(1 + E\sum_{\ell=0}^{\infty}(1+\epsilon_N)^\ell\, 1(\Gamma_{\ell+1} \le (2N+\theta)t)\,\psi_0(\ell) \times \left[1 + \frac{1}{\psi_0(N)}\sum_{i=0}^{\ell}(1 + \Gamma_{\ell+1} - \Gamma_{i+1})^{-d/2}\right]\Bigg) \quad (10.9)$$

It remains to show that the quantities in (10.5) and (10.9) approach 0 as $N \to \infty$. We begin by eliminating the contribution from large ℓ. Standard large deviations estimates for the sum of exponential mean one random variables (see Section 1.9 of Durrett (1995a)) imply that

Lemma 10.3. *If $A > 0$ is chosen large enough then for all p*

$$E\left(\sum_{\ell > AN} 1\{\Gamma_{\ell+1} \le (2N+\theta)t\} \cdot (1+\epsilon_N)^\ell\, \ell^p\right) \le C_p\, 2^{-N}$$

Using this result with the trivial fact that $m \le n$ implies $\Gamma_n - \Gamma_m \ge 0$, shows that the contributions to (10.5) and (10.9) from $\ell > AN$ approaches 0 as $N \to \infty$. To estimate the contributions from $\ell \le AN$, we begin with a simple estimate

Lemma 10.4. *If $p > 0$ there is a constant C_p so that if $m > p$ then*

$$E(1+\Gamma_m)^{-p} \le C_p(1+m)^{-p}$$

Proof. By a simple application of Jensen's inequality we may assume p is a positive integer. Clearly, $E(1+\Gamma_m)^{-p} \le E(\Gamma_m)^{-p}$. Integration shows that

$$E(\Gamma_m)^{-p} = \int_0^\infty x^{-p}\frac{x^{m-1}}{(m-1)!}e^{-x}\,dx = \frac{(m-1-p)!}{(m-1)!} \le C_p(1+m)^{-p}$$

□

Let (10.9b) denote the part of (10.9) that comes from $\ell \le AN$. Recall $\epsilon_N = \theta/(2N+\theta)$ and hence $(1+\epsilon_N)^{AN} \le C$. Using Lemma 10.4 now, and throwing away the indicator of $\Gamma_{\ell+1} \le (2N+\theta)t$, we see that (10.9b) is bounded by

$$\frac{C\tau_N}{\psi(N)} \cdot \{X_0^N(1)\} \cdot \psi_0(AN) \cdot \sum_{\ell=0}^{AN} \left[1 + \frac{1}{\psi_0(N)} \sum_{i=0}^{\ell} (1+\ell-i)^{-d/2} \right]$$

The quantity in square brackets is bounded and $\psi_0(AN) \le C\psi_0(N)$, so the above is at most $C\tau_N X_0^N(1)$ which approaches 0 as $N \to \infty$, and so (10.9) also approaches 0 as $N \to \infty$.

Let (10.5b) denote the part of (10.5) that comes from $\ell \le AN$. Recall that $\psi_0(N) = \psi(N)/N$. Discarding the indicator function of $\Gamma_{\ell+1} \le (2N+\theta)t$ as above, we may bound (10.5b) by

$$\frac{C}{N^2\psi_0(N)^2} \cdot X_0^N(1) \cdot E\left\{ \sum_{\ell=0}^{AN} \sum_{i=0}^{\ell} \sum_{k=-1}^{i-1} \sum_{j=0}^{\ell} 1(\Gamma_{\ell+1} - \Gamma_{i+1} \le (2N+\theta)\tau_N) \right.$$
$$\left. \times (1 + \Gamma_{i+1} - \Gamma_{k+1})^{-d/2} \left(1 + \Gamma_{\ell+1} - \Gamma_{j+1}\right)^{-d/2} \right\} \tag{10.10}$$

To attack this we will use the fact that if X is an indicator function (so $X^2 = X$) and Y and Z are nonnegative, then two applications of the Cauchy-Schwarz inequality imply

$$\begin{aligned} E(XYZ) &\le \{EX\}^{1/2}\{EY^2Z^2\}^{1/2} \\ &\le \{EX\}^{1/2}\{EY^4\}^{1/4}E\{Z^4\}^{1/4} \end{aligned}$$

This, together with Lemma 10.4, shows that (10.10) is at most

$$\frac{CX_0^N(1)}{N^2\psi_0(N)^2} \cdot \sum_{\ell=0}^{AN} \sum_{i=0}^{\ell} P(\Gamma_{\ell+1} - \Gamma_{i+1} \le (2N+\theta)\tau_N)^{1/2}$$
$$\times \sum_{k=-1}^{i-1} (1+i-k)^{-d/2} \sum_{j=0}^{\ell} (1+\ell-j)^{-d/2}$$

The sums over j and k are each smaller than $C\psi_0(AN) \le C'\psi_0(N)$, so the above is bounded by

$$\frac{CX_0^N(1)}{N^2} \cdot \sum_{\ell=0}^{AN} \sum_{i=0}^{\ell} P(\Gamma_{\ell-i} \le (2N+\theta)\tau_N)^{1/2} \tag{10.11}$$

To deal with this probability, note that a standard large deviations result for sums of exponentially distributed random variables (again see Section 1.9 of Durrett (1995a)) implies

Lemma 10.5. *There are constants* $0 < \gamma, C < \infty$ *so that if* $m \geq 2n$ *then*

$$P(\Gamma_m \leq n) \leq Ce^{-\gamma n}$$

From this it follows that if $\ell \leq AN$ then

$$\sum_{i=0}^{\ell} P(\Gamma_i \leq (2N+\theta)\tau_N)^{1/2} \leq 2(2N+\theta)\tau_N + AN \cdot Ce^{-\gamma(2N+\theta)\tau_N/2}$$

Using this in (10.11), then doing the sum over ℓ, which gives a factor of AN, we end up with an upper bound of

$$C\tau_N X_0^N(1)[\tau_N + e^{-\gamma(2N+\theta)\tau_N/2}] \to 0 \text{ as } N \to \infty \;.$$

This shows that (10.5) approaches 0 as $N \to \infty$ and so completes the proof of Lemma 10.1 and hence the proofs of Lemmas 6.3 and 6.7. This finishes our treatment of the interference term and hence the proof of our convergence theorem, Theorem 1.

11. Lower bound on the critical value

Let $\theta < b_d$ and let ξ_t denote the rescaled contact process starting from a single particle at the origin. Fix $t > 0$ and let Z_n be the discrete time branching random walk in which individuals in Z_{n-1} give birth to independent copies of ξ_t and hence multiple occupancy of sites is allowed. We view Z_n as an integer-valued measure on $\mathbf{R}^d$. It is easy to couple ξ_{nt} and Z_n so that Z_n dominates ξ_{nt}. Here note that particles in the contact processes underlying Z_n only have an offspring suppressed if they jump onto a site occupied by an offspring of the same parent in Z_{n-1} and there is no such ancestral restriction in the suppression of ξ offspring. If $E|\xi_t| < 1$, for N sufficiently large, then the subcritical Galton-Watson branching process $Z_n(1)$ dies out for large n $a.s.$ and so the same holds true for ξ_{nt}. Recall β_c is the critical value of β for which there is positive probability of survival as $t \to \infty$ for the contact process starting with a single occupied site. Thus, to prove the lower bound half of our asymptotics for the critical value in Theorem 2, it suffices to show

Lemma 11.1. *As* $N \to \infty$,

$$E\big(NX_t^N(1)\big|X_0^N = N^{-1}\delta_0\big) \to e^{(\theta-b_d)t} < 1$$

Proof. Let $X_t^N(i),\ i \leq N$ be $i.i.d.$ copies of X_t^N starting from $X_0^N = N^{-1}\delta_0$, and let $Y_t^N = \sum_{i=1}^N X_t^N(i)$. Then Y_t^N differs from X_t^N starting at δ_0 in that jumps onto an occupied site are suppressed only if the two colliding particles descended from the same ancestor at $t = 0$ and hence multiple occupancies are allowed. This in fact simplifies the proof of the main convergence result (Theorem 1) as Lemma 5.1 is no longer needed since $\beta_0 \neq \gamma_0$ is now incorporated into the killing term K_t. As this is the only place the non-atomic nature of X_0 is used we can drop this restriction, allow $Y_0 = \delta_0$ and conclude

$$Y_\cdot^N \text{ converges weakly to } X_\cdot\text{, super-Brownian starting at } \delta_0$$
$$\text{and with drift } \theta - b_d$$

This, combined with the fact that $E(Y_t^N(1)^2) \leq E[X_t^{0,N}(1)^2 | X_0^{0,N} = \delta_0]$ stays bounded as $N \to \infty$ (see Lemma 2.9), shows that

$$\lim_{N\to\infty} E\big(NX_t^N(1)\big|X_0^N = N^{-1}\delta_0\big) = \lim_{N\to\infty} E\big(Y_t^N(1)\big)$$
$$= E\big(X_t(1)\big) = e^{(\theta-b_d)t} \qquad \square$$

12. Upper bound on the critical value

Throughout this section we assume $\theta > b_d$. To prove the existence of a non-trivial stationary distribution and hence derive upper bounds on the critical value, we will use a rescaling argument to compare the long range contact process with oriented percolation. To establish the connection we begin by introducing the lattice on which percolation takes place:

$$\mathscr{L}_0 = \{(m, n) \in \mathbf{Z}^2 : m + n \text{ is even}, n \geq 0\}$$

Let $T > 0$ and $L = \sqrt{T} \in \mathbf{N}$. Let $I = [-L, L]^d$ be the cube of radius L, and let $e_1 = (1, 0, \ldots, 0)$ be the first unit vector. It will be convenient to assume $Le_1 \in \mathscr{Z}_N$ and so we will only consider N satisfying $N^{1/2+1/d} \in \mathbf{N}$ if $d \geq 3$, and for $d = 2$ replace $(\log N)^{1/2}$ by its integer part in the definition of $\mathscr{Z}_N$ and throughout the convergence theorem. This will ensure that $\mathbf{Z}^d \subset \mathscr{Z}_N$ and in particular $Le_1 \in \mathscr{Z}_N$.

Given a realization of the contact process ξ_t, and a site $(m, n) \in \mathscr{L}_0$ we will say that m is "occupied" at time n if the contact process when restricted to $I_m = 2Lme_1 + I$, and translated in space to be a function on I, lies in a set H of "happy" configurations. In words, the set will be chosen so that (i) if m is occupied at time n then with high probability $m + 1$ and $m - 1$ will be occupied at time $n + 1$, and (ii) the events that cause (i) to occur are determined by the behavior of the contact process modified so that particles which land outside of $2Lme_1 + (-KL, KL) \times \mathbf{R}^{d-1}$ are killed, for some fixed natural number K. The set H will be defined below but for now we note that the configuration which is 0 on I is not in H while the configuration of all 1's on I is.

More formally, we will check the comparison assumptions on p.140 of Section 4 of Durrett (1995b). Let $(\sigma_y\xi)(x) = \xi(x+y)$ denote the translation (or shift) of ξ by y and $\sigma_y H = \{\sigma_y\xi : \xi \in H\}$. For each $\gamma > 0$ and $K \in \mathbf{N}$ we introduce

$(CA)_{\gamma,K}$: For each $\xi \in H$ there is an event G_ξ, measurable with respect to the contact process with killing outside $(-KL, KL) \times R^{d-1} \times [0, T]$, and with $P(G_\xi) \geq 1 - \gamma$, so that on G_ξ, ξ_T lies in $\sigma_{2Le_1}H$ and in $\sigma_{-2Le_1}H$.

Here we consider rescaled ξ's which are therefore subsets of $\mathscr{Z}_N$, or equivalently $\{0, 1\}$-valued functions on $\mathscr{Z}_N$, and identify ξ with the measure $X^N(\xi)$ wich assigns mass $1/N$ to each site in ξ.

Legal scholars may have noted that page 140 of Durrett instead says "measurable with respect to the graphical representation," while in this paper we have used a branching process construction. However, it is easy to see that the construction used here has the property that if the space time boxes are disjoint then the subprocesses that result from the contact process restricted to these boxes are conditionally independent given their initial conditions. This is enough so that we can repeat the proof of Theorem 4.3 given in Durrett (1995b) in our new setting, and conclude that if

$$\chi_n = \{m : (m, n) \in \mathscr{L}_0, \xi_{nT} \in \sigma_{2Lme_1}H\},$$

then χ_n dominates an $2K$-dependent oriented percolation process (see (4.1) of Durrett (1995b)), W_n, with initial configuration $W_0 = \chi_0$ and density at least $1 - \gamma$, i.e., $W_n \subset \chi_n$ for all $n \geq 0$.

If, for a fixed value of K_0, we can check $(CA)_{\gamma,K_0}$ for all $\gamma > 0$ then taking $\xi_0^1(x) \equiv 1$ (which is in H and hence assures that χ_0 is the entire integer lattice) and using Theorem 4.2 in Durrett (1995b) gives

$$\liminf_{n\to\infty} P(0 \in \chi_n) > 0$$

From this and the fact that the configuration which is 0 on I is not in H it follows that the upper invariant measure $\bar{\xi}_\infty^1$ must be nontrivial. If not, then $\bar{\xi}_\infty^1 \equiv 0$ and

$$P(0 \in \chi_n) \leq P(\xi_{nT}(x) > 0 \text{ for some } x \in I) \to 0$$

Thus to complete the proof of the upper bound in Theorem 2 it suffices to check the comparison assumption.

Intuitively, to verify $(CA)_{\gamma,K_0}$ for the long range contact process, we will first verify $(CA)_{\gamma/2,K_0}$ for the limiting super-Brownian motion with drift $\theta - b_d > 0$ and then use our convergence theorem to conclude that if $\theta > b_d$ then $(CA)_{\gamma,K_0}$ holds for the contact process for large N. The set of configurations H that we will choose for the super-Brownian motion X, and for the rescaled contact processes X^N, are those that have enough mass and are not too concentrated. Specifically, $\mu \in H$ if there is a subconfiguration with corresponding measure $\nu \leq \mu$ with $\nu(I^c) = 0$, $\nu(I) = J_0$, and $Q(\nu) \leq q_0$, where Q is a quadratic form defined in (12.7) below, J_0 is a natural number selected in Choice 3 below and q_0 is a constant selected in Choice 4 below. Clearly H does not contain the configuration of all 0's. Moreover as we will be able to choose q_0 as large as we like and after the choice of J_0 (see Choice 4 below), it is clear that H will contain configuration of all 1's.

To check the comparison assumption we have to choose our constants to make the construction successful with high probability. To begin, we note that the limiting super-Brownian motion with drift $\theta - b_d$ has

$$E(X_T(I_1)|\, X_0 = \delta_x) = e^{(\theta - b_d)T} P_x(B_T \in I_1) \tag{12.1}$$

where B_t is a Brownian motion with variance $1/3$ per unit time. Easy calculations with the transition probability of Brownian motion show that

$$\liminf_{T\to\infty} \inf_{x\in I} P_x(B_T \in I_1) = \eta > 0 \tag{12.2}$$

This brings us to the first of several choices of parameters we will make.

Choice 1. *If $\theta > b_d$ we can pick $T \geq 1$ large enough so that $L = \sqrt{T} \in \mathbf{N}$ and*

$$\inf_{x\in I} E(X_T(I_1)|\, X_0 = \delta_x) \geq 5 \tag{12.3}$$

To achieve a finite range of dependence in our dominated percolation process, we need to impose a cutoff in space. In Bramson, Durrett and Swindle (1989) this was done by considering a modified contact process in which particles are killed if they move out of a finite strip. However, having worked for nine sections to prove the convergence of the rescaled contact process in the full space to super-Brownian motion, we do not want to repeat the proof for processes with killing outside of a strip, or ask the reader to believe we can do so. Thus we will take an approach that only requires use of the convergence theorem on the whole space. Let $\bar{X}_t^N$ be the Nth rescaled contact process modified so that particles born outside of $(-KL, KL) \times \mathbf{R}^{d-1}$ are immediately killed. Now the number of particles that are lost from the contact process by this truncation is at most the number that are lost in the dominating branching process $X_t^{0,N}$ (with "drift" θ).

The latter loss is easy to estimate. Let $\bar{X}_t^{0,N}$ be the Nth branching random walk, modified so that no particles are allowed to be born outside of

$(-KL, KL) \times \mathbf{R}^{d-1}$. For this it is well known, see, e.g., Sections 2 and 6 of Bramson, Durrett, and Swindle (1989), that the counterpart of Lemma 2.9 with killing. Namely

$$E\big(N\bar{X}_T^{0,N}(I_1)\big|\bar{X}_0^{0,N} = N^{-1}\delta_x\big) = e^{\theta T} P_x(\bar{B}_T^N \in I_1) \ , \tag{12.4}$$

where $\bar{B}_t^N$ is the random walk that takes jumps uniform on $\mathscr{N}_N$ at rate $N+\theta$, and is killed (i.e., sent to the state Δ) when it leaves $(-KL, KL) \times \mathbf{R}^{d-1}$. Using the L^2 maximal inequality on the first component of B_t^N it is easy to see that

Choice 2. *If* $K = K_0 \in \mathbf{N}^{\geq 2}$ *is large enough, then for all* $N \geq 1$

$$\sup_{x\in I} e^{\theta T} P_x(\bar{B}_T^N = \Delta) \leq 1 \tag{12.5}$$

Having fixed our time horizon T and our spatial truncation width K_0, our next step is to make the success probability high by using initial measures with large total mass. We do this both for our branching random walks $X^{0,N}$ and super-Brownian motion X.

Lemma 12.1. *There is a* $0 < C < \infty$ *so that for all natural numbers* J
(a) *If* $X_0^{0,N}(I) = J$ *and* $X_0^{0,N}(I^c) = 0$ *then*

$$P\big((X_T^{0,N} - \bar{X}_T^{0,N})(\mathbf{R}^d) \geq 2J\big) \leq C/J \ .$$

(b) *If* $X_0(I) \geq J$ *then* $P(X_T(I_1) \leq 4J) \leq C/J$.

Proof. (a) An easy calculation using (12.4) with $\mathbf{R}^d$ in place of I_1 and (12.5) shows that

$$E\big(\big(X_T^{0,N} - \bar{X}_T^{0,N}\big)(\mathbf{R}^d)\big) \leq X_0^{0,N}(I)$$

Turning to second moments, we have

$$\begin{aligned} &E(((X_T^{0,N} - \bar{X}_T^{0,N})(\mathbf{R}^d))^2) \\ &\quad = E[N^{-2} \sum_{\beta\sim T} \sum_{\gamma\sim T} 1(\zeta_\beta^0 > T, \zeta_\gamma^0 > T) \\ &\qquad \times 1(B_s^\beta, B_{s'}^\gamma \notin (-KL, KL) \times \mathbf{R}^{d-1} \text{ for some } s, s' \leq T)] \end{aligned}$$

The contribution from indices satisfying $\gamma_0 \neq \beta_0$ is at most $[E((X_T^{0,N} - \bar{X}_T^{0,N})(\mathbf{R}^d))]^2$. The contribution from indices satisfying $\beta_0 = \gamma_0$ is trivially bounded by

$$E\, N^{-2} \sum_{\beta\sim T} \sum_{\gamma\sim T} 1(\beta_0 = \gamma_0) 1(\zeta^0_\beta > T, \zeta^0_\gamma > T)$$
$$= X_0^{0,N}(\mathbf{R}^d) \left[N^{-1} e^{\theta T} + \left(1 + \frac{\theta}{N}\right) \theta^{-1} (e^{2\theta T} - e^{\theta T}) \right] \ ,$$

where we have used a well-known expression for the second moment of a branching random walk in the last line (see Lemma 2.2 of Bramson, Durrett and Swindle (1989)). The above calculations bound the variance of $(X_T^{0,N} - \bar{X}_T^{0,N})(\mathbf{R}^d)$ by $C(T) X_0^{0,N}(\mathbf{R}^d)$. The result now follows by Chebychev's inequality.

(b) This follows by a similar Chebychev argument using (12.3) to get a lower bound on the mean, and the fact that the variance of $X_T(I_1)$ is bounded by a constant times the initial mass (see Proposition (2.7) of Fitzsimmons (1988)). □

With Lemma 12.1 in mind and leaving lots of room for errors to accumulate, we can now make

Choice 3. *Let C be as in Lemma* 12.1, $\alpha = \gamma/100$ *and pick a natural number J_0 large enough so that $C/J_0 \le \alpha$.*

In order for the contact process to be successful at avoiding extinction with high probability, it is not sufficient that the initial number of particles is large. Consider for concreteness the situation in $d \ge 3$. In this case the neighborhood $\mathcal{N}_N$ has $O(N)$ particles. If we let the initial state consist of all the sites in one or more neighborhoods $x + \mathcal{N}_N$ then the mass lost due to births onto occupied sites will result in a devastating decrease. Since we will not need to know the details, we leave it to the reader to figure out how much mass is lost and how quickly. To avoid this problem, we let

$$\ell(z) = \begin{cases} \log(1/\|z\|_\infty) & \text{for } 0 < \|z\|_\infty < 1 \\ 0 & \text{if } z = 0 \text{ or } \|z\|_\infty \ge 1 \end{cases} \tag{12.6}$$

define the quadratic form

$$Q(\mu) = \iint \mu(dx)\, \mu(dy)\, \ell(y - x) \tag{12.7}$$

and then consider initial conditions for the contact process that are supported in I, have $X_0^N(I) = J_0$, $Q(X_0^N) \le M$, and, of course, at most one particle per site. Note that we have set $\ell(0) = 0$ to avoid the infinities on the diagonal $x = y$ when we are dealing with point mass measures.

Using our convergence theorem now with Lemma 12.1 and our choice of J_0 we have

Lemma 12.2. *Let $M \geq 1$. If $N \geq N_0(M)$, $X_0^N(I^c) = 0$, $X_0^N(I) = J_0$, and $Q(X_0^N) \leq M$ then*

$$P(X_T^N(I_1) \leq 4J_0) \leq 2\alpha$$

Proof. If not, then there is a subsequence of integers $N_k \uparrow \infty$ and associated initial conditions $X_0^{N_k}$ where the probability exceeds 2α. Since the measures $X_0^{N_k}$ have support in I and total mass J_0 there is a weakly convergent subsequence. The limit must be atomless by Fatou's lemma, the bound on $Q(X_0^N)$, and the lower semicontinuity of ℓ. Our convergence theorem shows that (recall I_1 is defined to be open)

$$P\Big(X_T(I_1) \leq 4J_0\Big) \geq \limsup_{k\to\infty} P\Big(X_T^{N_k}(I_1) \leq 4J_0\Big) \geq 2\alpha$$

which contradicts (b) of Lemma 12.1 and the choice of J_0. □

Lemma 12.3. *Let $M \geq 1$. If $N \geq N_1(M)$ then for all X_0^N with $X_0^N(I^c) = 0$, $X_0^N(I) = J_0$, and $Q(X_0^N) \leq M$ we have*

$$P(\bar{X}_T^N(I_1) \leq 2J_0) \leq 3\alpha$$

Proof. As noted above we can bound the amount of mass lost in the contact process by the mass lost in the branching process, so (a) of Lemma 12.1 implies

$$P\left((X_T^N - \bar{X}_T^N)(I_1) \geq 2J_0\right) \leq C/J_0$$

The desired result now follows from Lemma 12.2 and the choice of J_0. □

Having imposed the condition $Q(X_0^N) \leq M$ on the initial condition, we are now obliged to show that with high probability it holds at time T. To do this it is enough to show the following result for the dominating branching random walks $X_t^{0,N}$.

Lemma 12.4. *For any natural number J there is a $0 < C_{T,J} < \infty$ and N_2 so that if $N \geq N_2$ then for all $X_0^{0,N}$ with $X_0^{0,N}(\mathbf{R}^d) \leq J$ we have*

$$EQ(X_T^{0,N}) \leq C_{T,J}$$

This should motivate the final

Choice 4. *Pick $q_0 \geq 1$ large enough so that $C_{T,J_0}/q_0 \leq \alpha$.*

To verify $(CA)_{\gamma,K_0}$, for $\xi \in H$ choose $\nu \leq \xi$ as in the definition of H (considering ξ as a measure) and let G_ξ be the event that, starting with ν, our modified contact process with killing, $\bar{X}^N$, satisfies $Q(\bar{X}_T^N) \leq q_0$, $\bar{X}_T^N(I_1) \geq J_0$, and $\bar{X}_T^N(I_{-1}) \geq J_0$. Here we choose $N \geq N_1(q_0) \vee N_2(J_0)$ so that Lemmas 12.3 and 12.4 are available with $M = q_0$ and $J = J_0$, respectively. This modified contact process uses the same exponential variables to jump or die as the full contact process X^N starting from ξ and so it is readily seen that the modified process is dominated by X^N. By using the collection of sites in $\bar{X}_T^N$ as our choice of $\nu \leq X_T^N$, we therefore see that $\xi_T \in \sigma_{2Le_1} H \cup \sigma_{-2Le_1} H$. Finally Lemmas 12.3 and 12.4 show that

$$\begin{aligned} P(G_\xi^c) &\leq P(\bar{X}_T^N(I_1) \leq 2J_0 | \bar{X}_0^N = \nu) + P(\bar{X}_T^N(I_{-1}) \leq 2J_0 | \bar{X}_0^N = \nu) \\ &\quad + E(Q(\bar{X}_T^N) | \bar{X}_0^N = \nu)/q_0 \\ &\leq 6\alpha + C_{T,J_0}/q_0 \leq 7\alpha < \gamma \end{aligned}$$

and so $(CA)_{\gamma,K_0}$ holds. Thus the last detail is to complete the

Proof of Lemma 12.4. Using $\beta \approx T$ as shorthand for $T_{\pi\beta} < T < T_\beta$ and $B^\beta \neq \Delta$, we have

$$N^2 \cdot EQ(X_T^{0,N}) = E\left(\sum_{\beta,\gamma} 1(\beta \approx T, \gamma \approx T) \cdot \ell(B^\gamma - B^\beta)\right) \tag{12.8}$$

First consider the β and γ with $\beta_0 = i$ and $\gamma_0 = j$ where $i \neq j$. Imitating (4.4)–(4.6) we can write (recall $\epsilon_N = \theta(2N+\theta)^{-1}$)

$$\begin{aligned} &\sum_{\beta:\beta_0=i} \sum_{\gamma:\gamma_0=j} E\big(1(\beta \approx T, \gamma \approx T) \cdot \ell(B^\gamma - B^\beta)\big) \\ &\quad = \sum_{n=0}^{\infty} (1+\epsilon_N)^n e^{-(4N+2\theta)T} \frac{((4N+2\theta)T)^n}{n!} E\ell\left(x_j - x_i + N^{-1/2} V_n^N\right) \end{aligned} \tag{12.9}$$

where V_n^N is the random walk that stays put with probability 1/2 and with probability 1/2 takes a jump uniformly distributed over $N^{1/2}\mathcal{N}_N$.

For all $z \in \mathscr{L}_N$ we have $\ell(z) \leq C \log N$ so large deviations results for the Poisson distribution imply

$$(C \log N) \cdot \sum_{n=2(4N+2\theta)T}^{\infty} (1+\epsilon_N)^n \cdot e^{-(4N+2\theta)T} \frac{((4N+2\theta)T)^n}{n!} \to 0 \quad \text{as } N \to \infty \tag{12.10}$$

For the sum over $n < 2(4N + 2\theta)T$ note that $(1 + \epsilon_N)^n \le C_T$, so that part is bounded by

$$C_T \sum_{n=0}^{\infty} e^{-(4N+2\theta)T} \frac{((4N + 2\theta)T)^n}{n!} E\ell(x_j - x_i + N^{-1/2} V_n^N)$$
$$= C_T E\ell\left(x_j - x_i + N^{-1/2} V^N_{\pi((4N+2\theta)T)}\right)$$

where $\pi(u)$, $u \ge 0$ is a Poisson process with rate one.

Let T_1 be the time of the first jump of $\pi(u)$. Using (8.7) we can estimate (recall $T \ge 1$)

$$P\big(T_1 \le T(2N + \theta),\, V^N_{\pi((4N+2\theta)T)} - V_1^N \in x + [-1, 1]^d\big) \le CN^{-d/2}$$

Since $P(T_1 > T(2N+\theta)) \le e^{-(2N+\theta)T}$, considering value of the first jump shows

$$P(N^{-1/2} V^N_{\pi((4N+2\theta)T)} = x) \le \frac{C}{N^{d/2}\psi(N)} \tag{12.11}$$

Since ℓ is constant on $\{x : \|x\|_\infty = c\}$ and decreasing for $0 < c < 1$, the maximum value of $E\ell(x_j - x_i + N^{-1/2} V^N_{\pi((4N+2\theta)T)})$ subject to the constraint on the probabilities in (12.11) can be bounded by

$$C \iint \ell(y - x) d\mu_N(x) d\mu_N(y)$$

where μ_N is the uniform distribution on the points of $\mathscr{L}_N$ in $[-1, 1]^d$. This easily gives

$$E\ell\left(x_j - x_i + N^{-1/2} V^N_{\pi((4N+2\theta)T)}\right) \le C \int_{[-1,1]^{2d}} \ell(y - x) dx\, dy \tag{12.12}$$

We note that as usual the value of C changes from line to line in the above. Summing over i and j now gives

$$E\left(\sum_{\beta,\gamma:\beta_0\ne\gamma_0} 1(\beta \approx T, \gamma \approx T) \cdot \ell(B^\gamma - B^\beta)\right) \le C_T(1 + \{NX_0^{0,N}(1)\}^2) \tag{12.13}$$

Turning to the terms with $\beta_0 = \gamma_0$, we let $\alpha = \beta \wedge \gamma$ and note that as $\ell(0) = 0$, we take $\beta \ne \gamma$ and so $k = |\alpha| < |\beta| \wedge |\gamma|$ in the above sum. Let $\ell \ge 1$ be such that $|\beta| = k + \ell$ and $m \ge 1$ be such that $|\gamma| = k + m$. Arguing as in Lemma 4.4, we have

$$E\left(\sum_{\substack{\beta,\gamma:\beta\wedge\gamma=\alpha \\ |\beta|=k+\ell,|\gamma|=k+m}} 1(\beta \approx T, \gamma \approx T)\ell(B^\gamma - B^\beta) \middle| \mathscr{H}_\alpha \right)$$
$$\leq (1+\epsilon_N)^{\ell-1+m-1} e^{-2(2N+\theta)(T-T_\alpha)} \frac{((2N+\theta)(T-T_\alpha))^{\ell-1+m-1}}{(\ell-1)!(m-1)!}$$
$$\times \left[E\ell\left(N^{-1/2} V^N_{\ell+m-1}\right)\right] 1(T_\alpha \leq T, T_\alpha \leq \zeta^0_\alpha)$$

Summing over the possible values of ℓ and m, changing variables $n = \ell + m - 2$, and noting

$$\sum_{m=1}^{n+1} \frac{1}{(n-m+1)!(m-1)!} = \frac{2^n}{n!} \ ,$$

we have

$$E\left(\sum_{\beta,\gamma:\beta\wedge\gamma=\alpha} 1(\beta \approx T, \gamma \approx T)\ell(B^\gamma - B^\beta) \middle| \mathscr{H}_\alpha \right)$$
$$\leq \sum_{n=0}^{\infty} (1+\epsilon_N)^n e^{-(4N+2\theta)(T-T_\alpha)} \frac{((4N+2\theta)(T-T_\alpha))^n}{n!}$$
$$\times \left[E\ell\left(N^{-1/2} V^N_{n+1}\right)\right] 1(T_\alpha \leq T, T_\alpha \leq \zeta^0_\alpha)$$

As in (12.10), large deviations results for the Poisson distribution imply that the sum over $n \geq 2(4N+2\theta)T$ is bounded by $Ce^{-cN} 1(T_\alpha \leq T, T_\alpha \leq \zeta^0_\alpha)$ for some $c > 0$. Continuing to reason as in the case $\beta_0 \neq \gamma_0$ we see that the sum over $n \leq 2(4N+2\theta)T$ is bounded by

$$C_T\, E\ell\left(N^{-1/2} V^N_{1+\pi((4N+2\theta)(T-T_\alpha))}\right) \equiv C_T E\ell\left(\bar{V}^N(T-T_\alpha)\right)$$

Repeating the proof of (12.11) now shows that

$$P\Big(\bar{V}^N(T-T_\alpha) = x\Big) \leq \frac{C}{\psi(N)} \cdot \Big(1 + (4N+2\theta)(T-T_\alpha)\Big)^{-d/2}$$

Again maximizing with respect to this constraint on the probabilities, gives

$$E\ell\left(\bar{V}^N(T-T_\alpha)\right)$$
$$\leq C \sum_{x\in\mathscr{L}_{\mathscr{N}}} 1(\|x\|_\infty \leq (T-T_\alpha)^{1/2})\ell(x)\psi(N)^{-1}$$

$$\times(1+4(N+2\theta)(T-T_\alpha))^{-d/2}+\ell((T-T_\alpha)^{1/2})$$
$$\le C(1+\ell(T-T_\alpha))$$

Summing over α now gives

$$E\left(\sum_{\beta,\gamma:\beta_0=\gamma_0} 1(\beta\approx T,\gamma\approx T)\ell(B^\gamma-B^\beta)\right)$$
$$\le C_T\left(1+E\left(\sum_\alpha 1(T_\alpha\le T, T_\alpha\le \zeta^0_\alpha)(1+\ell(T-T_\alpha))\right)\right)$$

It follows from Lemma 3.4(b) that the above equals

$$C_T\left(1+\int_0^T e^{\theta s}NX_0^{0,N}(1)\cdot(1+\ell(T-s))\cdot(2N+\theta)ds\right)$$
$$\le C_T(1+N^2X_0^{0,N}(1))$$

Combining this with (12.13), we have the desired bound on (12.8). This completes the proof of Lemma 12.4. □

References

Billingsley, P.: Convergence of Probability Measures. John Wiley, New York, (1968)

Bramson, M., Durrett, R., Swindle, G.: Statistical mechanics of crab grass. Ann. Prob., **17**, 444–481 (1989)

Breiman, L.: Probability. Addison-Wesley, Reading, Mass (1968)

Buttel L., Cox, T., Durrett, R.: Estimating the critical values of stochastic growth models. J. Appl. Prob., **30**, 445–461 (1993)

Dawson, D.A.: Geostochastic calculus. Can. J. Statistics, **6**, 143–168 (1978)

Dawson, D.A.: Measure-valued Markov processes. Pages 1–260 in Ecole d'été de Probabilités de St. Flour, XXI Lecture Notes in Math 1541, Springer, New York (1993)

Durrett, R.: Lecture Notes on Particle Systems and Percolation. [Chapters 4 and 11] Wadsworth Pub. Co. (1988)

Durrett, R.: A new method for proving the existence of phase transitions. Pages 141–170 in Spatial Stochastic Processes Edited by K.S. Alexander and J.C. Watkins. Birkhauser, Boston (1992a)

Durrett, R.: The contact process: 1974–89. Pages 1–18 in Mathematics of Random Media Edited by W. Kohler and B. White. American Math. Society, Providence, RI (1992b)

Durrett, R.: Probability: Theory and Examples. Duxbury Press (1995a)

Durrett, R.: Ten lectures on particle systems. Pages 97–201 in Ecole d'été de Probabilités de St. Flour, XXIII Lecture Notes in Math 1608, Springer, New York (1995b)

Ethier, S., Kurtz, T.: Markov Processes: Characterization and Convergence. John Wiley, New York (1986)

Evans, S.N., Perkins, E.: Measure-valued branching diffusions with singular interactions. Can. J. Math., **46**, 120–168 (1994)

Fitzsimmons, P.J.: Construction and regularity of measure-valued Markov branching processes. Israel J. Math., **64**, 337–361 (1988)

Harris, T.: Contact interactions on a lattice. Ann. Prob., **2**, 969–988 (1974)

Holley, R., Liggett, T.: The survival of contact processes. Ann. Prob., **6**, 198–206 (1978)

Kendall, D.: Stochastic processes and population growth. J. of the Roy. Stat. Soc., **11**, 230–264 (1949)

Kesten, H.: A sharper form of the Doeblin-Levy-Kolmogorov-Rogozin inequality for concentration functions. Math. Scand., **25**, 133–144 (1969)

Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer-Verlag, Berlin (1987)

Konno, N., Shiga, T.: Stochastic differential equations for some measure-valued diffusions. Prob. Theory Rel. Fields, **79**, 201–225 (1988)

Liggett, T.: Interacting Particle Systems. Springer-Verlag, New York (1985)

Liggett, T.: Improved upper bounds for the contact process critical value. Ann. Prob., **23**, 697–723 (1995)

Mueller, C., Tribe, R.: A phase transition for a stochastic PDE related to the contact process. Prob. Theory Rel. Fields, **100**, 131–156 (1994)

Mueller, C., Tribe, R.: Stochastic PDE's arising from the long range contact process and long range voter model. Prob. Theory Rel. Fields, **102**, 519–546 (1995)

Reimers, M.: One dimensional stochastic partial differential equations and the branching measure diffusion. Prob. Theory Rel. Fields, **81**, 319–340 (1989)

Stacey, A.: Bounds on the critical probabilities in oriented percolation models, Ph.D. Thesis, Cambridge U (1994)

Walsh, J.: An introduction to stochastic partial differential equations. Pages 265–439 in Ecole d'été de Probabilités de St. Flour, XIV Lecture Notes in Math 1180, Springer, New York (1986)